

# **PROSPECTIVE SENSE OF AGENCY: COGNITIVE AND NEURAL MECHANISMS**

**Nura Cabral Sidarus**

Institute of Cognitive Neuroscience

University College London

PhD Thesis

2016



## **Declaration**

I, Nura Cabral Sidarus, confirm that the work presented in this thesis is my own. Any information derived from other sources is fully cited and referenced in the thesis.

Nura Cabral Sidarus





## **Abstract**

The sense of agency (SoA) refers to the sense of being in control of one's actions and, through them, of events in the external world. Much research has focused on how we link our actions to their outcomes. However, the contribution to SoA of processes linking our intentions to our actions has received little attention. The present research focused, therefore, on investigating the cognitive and neural mechanisms through which action selection processes can prospectively inform our SoA. Recent work revealed that influencing action selection through subliminal priming can lead to a reduction in SoA. Here, new tasks and manipulations of action selection were developed. The generalizability of previous results was demonstrated – a consistent reduction in SoA was found when action selection was disrupted. This effect was found for: disruptions at different stages of action selection; for different levels of awareness of distracting stimuli; and were unaffected by whether participants freely chose what to do, or followed an instruction. In other experiments, SoA judgements were tested in the context of a computer game, providing a more ecological and dynamic context than previous studies.

Electrophysiological investigations of the neural correlates of agency showed neural monitoring of the action itself was reliably associated with judgements of agency, independently from and in addition to previously-established neural processes for monitoring outcomes. These findings, together with a meta-analysis of available studies, support a dissociation between prospective and retrospective components of SoA. Finally, the influence of social context on action selection and SoA was explored. Under conditions in which outcomes were unambiguously self-caused, the presence of an alternative agent who could act instead of oneself led to both reduced SoA and attenuated neural processing of those outcomes. The prospective sense of agency may be important as an advance predictor of successful action, allowing for immediate corrective action, as well as for learning to adapt behaviour in the future.



## Acknowledgements

I am immensely grateful to my supervisor, Patrick Haggard, for the amazing journey of the last few years. It has been a great privilege to work with, and learn from, such a creative and inspiring mentor. I have learned so much from our meetings, not only about experimental work, but also about how to organise and improve one's thinking, and how to communicate one's work and ideas effectively. His trust and guidance throughout have been invaluable. I also have enormous appreciation for the opportunity to work in such a rich and diverse lab. I am very thankful to the many members (too many to list) who helped create a highly enjoyable and stimulating environment. In particular, I would like to thank Frederike Beyer for her vital collaboration for the study in Chapter 6; and to Jack de Havas, Brianna Beck and Steven di Costa for their time and thoughtful comments on previous versions of this thesis.

I would also like to thank Janet Metcalfe for the opportunity to spend some very enjoyable and instructive months working in her lab, at Columbia University. I especially want to thank Matti Vuorre for all his help setting up experiments during my stay, and with data analyses, in addition to our many stimulating discussions.

Finally, I am very grateful to my family for their endless support and encouragement. In particular, my sister Muna for the countless and very helpful discussions on the work in this thesis. I would also like to thank my friends Inke, Luís and Paulina, for their support, helpful discussions, and being willing victims of practice presentations, pilot testing, or proofreading for this thesis.



## Contributions

Some of the work presented in this thesis was conducted in collaboration with other researchers. I am very grateful for the contributions and help of these researchers and students.

The data collection for the experiments described in Chapter 2 was completed with the help of Ksenia Vinogradova.

The ideas for the experiments in Chapter 3 were developed with Patrick Haggard. The experiments were conducted at Columbia University, under the supervision of Janet Metcalfe. Matti Vuorre assisted with the programming and analysis. Data collection was completed with the help of lab research assistants Nina Plotnikov and Luigia Goodman.

Data collection for the experiment in Chapter 4 counted with the help of Ksenia Vinogradova. Data analyses were assisted by Matti Vuorre, who also reviewed the manuscript.

The meta-analysis described in Chapter 5 was done in collaboration with Matti Vuorre, who conducted the analyses, and contributed to the manuscript.

The study in Chapter 6 was done in collaboration with Frederike Beyer and Sofia Bonicalzi. The experimental design was developed in conjunction with Frederike Beyer, who also completed part of the data collection, analyses, and writing of the manuscript. Sofia Bonicalzi assisted with the data collection, and with reviewing the manuscript.



# Contents

Chapter 1. General Introduction.....	19
1.1. Introduction.....	19
1.1.1. Why is it important? .....	20
1.1.2. Summary .....	22
1.1.3. Motivation for this research.....	22
1.2. Research on Agency .....	24
1.2.1. How can we measure the sense of agency? .....	24
1.2.2. How does the sense of agency come about?.....	26
1.3. Sense of Agency and Other Mental Functions.....	36
1.3.1. Metacognition .....	36
1.3.2. Social Aspects of Agency .....	40
1.4. Limitations of the Previous Literature .....	41
1.4.1. Focus on retrospective agency attribution .....	41
1.4.2. Current prospective models remain unclear .....	42
1.4.3. Does agency reflect fluency or effort? .....	42
1.5. Aims of the Present Thesis .....	43
Chapter 2. Difficult Action Decisions Reduce the Sense of Agency: A Study Using the Eriksen Flanker Task .....	47
2.1. Introduction.....	48
2.2. Experiment 1 .....	53
2.2.1. Materials and Methods .....	53
2.2.2. Results.....	57
2.2.3. Discussion .....	59
2.3. Experiment 2 .....	62
2.3.1. Materials and Methods .....	62
2.3.2. Results.....	63
2.3.3. Discussion .....	66
2.4. Experiment 3 .....	67
2.4.1. Materials and Methods .....	67
2.4.2. Results.....	69
2.4.3. Discussion .....	71
2.5. General Discussion .....	73
2.5.1. Conclusions .....	78
Chapter 3. New Avenues for Investigating the Prospective Sense of Agency.....	81
3.1. Introduction.....	82
3.2. Experiment 1 .....	85
3.2.1. Materials and Methods .....	85
3.2.2. Results.....	88
3.2.3. Discussion .....	90
3.3. Experiment 2 .....	92

3.3.1. Materials and Methods .....	92
3.3.2. Results.....	95
3.3.3. Discussion .....	97
3.4. Experiment 3 .....	100
3.4.1. Materials and Methods .....	100
3.4.2. Results.....	102
3.4.3. Discussion .....	105
3.5. General Discussion .....	107
Chapter 4. How Action Selection Influences the Sense of Agency: an ERP study .....	113
4.1. Introduction.....	114
4.1.1. Action Monitoring .....	115
4.1.2. Outcome monitoring .....	115
4.1.3. Present study.....	116
4.2. Materials and Methods.....	117
4.2.1. Participants.....	117
4.2.2. Apparatus and Materials.....	117
4.2.3. Design and Procedure .....	118
4.2.4. EEG Recording and Analysis .....	121
4.3. Results .....	123
4.3.1. Agency Task.....	123
4.3.2. Prime Awareness Test.....	126
4.3.3. ERPs.....	126
4.4. Discussion .....	135
4.4.1. Action Monitoring .....	135
4.4.2. Outcome monitoring .....	140
4.4.3. Common performance monitoring framework .....	143
4.4.4. Conclusions .....	145
Chapter 5. Integrating Prospective and Retrospective Cues to the Sense of Agency: A Multi-Study Investigation.....	147
5.1. Introduction.....	148
5.2. Methods.....	152
5.2.1. Experimental Design.....	152
5.2.2. Study Selection for Multi-Study Analysis .....	152
5.2.3. Data Analysis.....	153
5.3. Results .....	155
5.3.1. Multilevel Model of Agency Ratings.....	156
5.4. Discussion .....	159
5.4.1. Prospective vs. Retrospective Cues .....	160
5.4.2. Competition Between Retrospective Cues .....	163
5.4.3. Conclusions .....	165



Chapter 6. Beyond Self-serving Bias: Diffusion of Responsibility Reduces Sense of Agency and Outcome Monitoring .....	167
6.1. Introduction.....	168
6.2. Materials and Methods .....	171
6.2.1. Participants .....	171
6.2.2. Apparatus and Materials.....	172
6.2.3. Design and Procedure .....	172
6.3. ERP Preprocessing .....	175
6.4. Data analysis.....	176
6.5. Results .....	176
6.5.1. Behaviour.....	176
6.5.2. ERPs.....	179
6.6. Discussion .....	181
6.6.1. Behaviour.....	181
6.6.2. FRN .....	182
6.6.3. Implications for concepts of diffusion of responsibility .....	183
6.6.4. Limitations and outlook.....	185
6.6.5. Conclusions .....	186
Chapter 7. General Discussion .....	189
7.1. Prospective Sense of Agency .....	190
7.1.1. Effects of Disrupted Action Selection.....	190
7.1.2. SoA and Post-decisional Action Monitoring.....	192
7.2. Integrating Prospective and Other Cues to Sense of Agency .....	194
7.2.1. Role of Awareness of Biases and Choice.....	195
7.2.2. Lack of Relation to Outcome Predictions.....	195
7.2.3. Interactions with Outcome Monitoring .....	196
7.3. Varieties of Agency Experience .....	198
7.4. Implications for Responsibility .....	200
7.5. Remaining Issues.....	201
7.5.1. Fluency vs. Effort.....	201
7.5.2. Control over Actions vs. Control over Outcomes.....	202
7.5.3. What is Prospective Agency for?.....	203
Appendix A. Supplementary Analyses for Chapter 2.....	205
Appendix B. Supplementary Analyses for Chapter 3.....	207
B.1. Experiment 1 .....	209
B.2. Experiment 2.....	210
B.3. Experiment 3.....	212
Appendix C. Supplementary Data for Chapter 4.....	213
C.1. Target N2.....	213
C.2. Action CRN .....	213
C.2.1. Action CRN model .....	213

C.2.2. Action CRN model with Target N2 .....	214
C.3. Outcome FRN .....	214
C.3.1. Outcome FRN model .....	214
C.3.2. Outcome FRN model with Action CRN .....	215
C.4. Outcome N1 .....	215
Appendix D. Supplementary Analyses for Chapter 5 .....	216
D.1. Group 1 .....	216
D.2. Group 2 .....	218
References	220

## List of Figures

<b>Figure 1.1.</b> Two levels of the sense of agency .....	25
<b>Figure 1.2.</b> Schematic of the computational model for sensorimotor control .....	30
<b>Figure 2.1.</b> Timeline of an example incongruent trial, with trial-wise ratings. ....	55
<b>Figure 2.2.</b> Results for Experiment 1 .....	58
<b>Figure 2.3.</b> Results for Experiment 2 .....	65
<b>Figure 2.4.</b> Results for Experiment 3 .....	70
<b>Figure 2.5.</b> Prospective and retrospective contributions to the sense of agency .....	77
<b>Figure 3.1.</b> Task outline for Experiment 1 .....	87
<b>Figure 3.2.</b> Results of Experiment 1 .....	89
<b>Figure 3.3.</b> JoAs model for Experiment 1 .....	90
<b>Figure 3.4.</b> Schematic of stimuli in Experiment 2 .....	94
<b>Figure 3.5.</b> Results of Experiment 2 .....	95
<b>Figure 3.6.</b> JoAs Model for Experiment 2 .....	97
<b>Figure 3.7.</b> Schematic of stimuli in Experiment 3 .....	102
<b>Figure 3.8.</b> Results of Experiment 3 .....	103
<b>Figure 3.9.</b> JoAs Model for Experiment 3 .....	104
<b>Figure 4.1.</b> Experimental paradigm. ....	119
<b>Figure 4.2.</b> Behavioural results .....	125
<b>Figure 4.3.</b> Target-locked ERPs across choice and priming conditions. ....	127
<b>Figure 4.4.</b> Target N2 model predictions for the priming by RTs interaction .....	128
<b>Figure 4.5.</b> Action CRN model predictions for the priming by RTs interaction .....	129

<b>Figure 4.6.</b> Outcome-locked ERPs across choice and priming conditions. ....	131
<b>Figure 4.7.</b> Parameter estimates for the model predicting agency ratings .....	133
<b>Figure 4.8.</b> Relation between subjective agency ratings and ERPs .....	134
<b>Figure 4.9.</b> Schematic outline of the experimental paradigm with the putative cognitive processes involved at different stages of voluntary action. ....	136
<b>Figure 5.1.</b> Rationale for the study. ....	151
<b>Figure 5.2.</b> Task outline for the subliminal priming studies. ....	153
<b>Figure 5.3.</b> Mean agency rating across participants as a function of trial number, for each experiment .....	155
<b>Figure 5.4.</b> Mean agency rating as function of AOI for each experiment .....	156
<b>Figure 5.5.</b> Multilevel model of agency ratings. ....	159
<b>Figure 6.1.</b> Marble Task.....	173
<b>Figure 6.2.</b> Behavioural results.....	177
<b>Figure 6.3.</b> Outcome-locked ERPs. ....	180
<b>Figure 7.1.</b> Four distinctive varieties of agency experience.....	199

## List of Tables

<b>Table 5.1.</b> Factorial design and differences between studies. ....	154
<b>Table 5.2.</b> Multilevel model of agency ratings.....	158
<b>Table 6.1.</b> Agency ratings model .....	178
<b>Table 6.2.</b> FRN model: parameter estimates .....	180

## List of Appendix Figures

<b>Figure B1.</b> Experiment 1 .....	208
<b>Figure B2.</b> Experiment 2.....	208
<b>Figure B3.</b> Experiment 3.....	208

## List of Appendix Tables

<b>Table A1.</b> Model for Experiment 2 predicting agency ratings .....	205
<b>Table A2.</b> Model for Experiment 3 predicting agency ratings .....	205

<b>Table B1.</b> Experiment 1, Model 1 – Linear effect of JoPs.....	209
<b>Table B2.</b> Experiment 1, Model 2 – Quadratic effect of JoPs. ....	210
<b>Table B3.</b> Experiment 1, Model 3 – Cubic effect of JoPs.....	210
<b>Table B4.</b> Experiment 1: Model comparison.....	210
<b>Table B5.</b> Experiment 2, Model 1 – Linear effect of JoPs.....	210
<b>Table B6.</b> Experiment 2, Model 2 – Quadratic effect of JoPs. ....	211
<b>Table B7.</b> Experiment 2, Model 3 – Cubic effect of JoPs.....	211
<b>Table B8.</b> Experiment 2: Model comparison.....	211
<b>Table B9.</b> Experiment 3, Model 1 – Linear effect of JoPs.....	212
<b>Table B10.</b> Experiment 3, Model 2 – Quadratic effect of JoPs. ....	212
<b>Table B11.</b> Experiment 3, Model 3 – Cubic effect of JoPs.....	212
<b>Table B12.</b> Experiment 3: Model comparison.....	212
<b>Table C1.</b> Target N2 model.....	213
<b>Table C2.</b> Action CRN model .....	213
<b>Table C3.</b> Action CRN model with Target N2.....	214
<b>Table C4.</b> Outcome FRN model.....	214
<b>Table C5.</b> Outcome FRN model with Action CRN.....	215
<b>Table C6.</b> Outcome N1 model.....	215
<b>Table D1.</b> Multilevel model of agency ratings for Experiment 1.....	217
<b>Table D2.</b> Multilevel model of agency ratings for Experiment 2.....	217
<b>Table D3.</b> Multilevel model of agency ratings for Experiment 3.....	218
<b>Table D4.</b> Multilevel model of agency ratings for Experiment 4.....	218
<b>Table D5.</b> Multilevel model of agency ratings for Experiment 5.....	219
<b>Table D6.</b> Multilevel model of agency ratings for Experiment 6.....	219
<b>Table D7.</b> Multilevel model of agency ratings for Experiment 7.....	219





# Chapter 1. General Introduction

## 1.1. Introduction

The sense of being an agent in the world, of controlling one's actions and, through those actions, controlling events in the external world, is a fundamental aspect of human experience. This sense of agency (SoA) underpins our motivation for goal-directed actions (Bandura, 1991), and thus our uniquely human capacity to radically change the world around us. The phenomenal experience of being an active agent is also central to our sense of self (Knoblich, Elsner, Aschersleben, & Metzinger, 2003), as it supports the distinction between our bodies and ourselves, and the outside world. The sense of agency typically colours the background of our mental lives, but we become acutely aware of it when the smooth flow of voluntary action, from intention, to action, to outcome, is disrupted (Chambon, Sidarus, & Haggard, 2014; Gallagher, 2012).

For example, I may walk into a dark room while talking to a friend, and flip the light switch. If the lights turn on, as expected, I may barely notice that I flipped the switch. However, if the lights fail to turn on, I will suddenly become very aware of my action and its consequences (or lack thereof). Even when the sense of agency is not at the forefront of our conscious experience, our brain is constantly monitoring our actions and their consequences, capturing our attention whenever anything goes awry. At the same time, we are able to intentionally introspect on our experience of agency. If my friend were to stop me while I was flipping the light switch and ask what I was doing, I would quickly reply that I was turning on the lights. SoA has a pre-reflective, experiential component, as well as a more reflective, evaluative component (Gallagher, 2012; Synofzik, Vosgerau, & Newen, 2008).

### 1.1.1. Why is it important?

#### ***Learning vs. Expert Control of Behaviour.***

These different aspects of SoA are also salient when learning a new skill. When we start learning a new skill, like driving, we are acutely aware of all our movements, and of tracking the link between our actions and their consequences. Here, our experience of agency, and of learning how to control our environment is central to our conscious experience, even when all goes according to plan. As we become experts at the task, performing the various necessary actions, e.g. changing gears or pressing pedals, becomes more automatic, more fluent, and our SoA tends to be less salient. Similarly, expertise can be associated with an experience of “flow” (Csikszentmihalyi, 1990), in which one is fully immersed in the task at hand, and in perfect control over one’s actions and their consequences. Paradoxically, as a result of the strong focus on the task, some people report a loss of self in flow states, which leads them to feel *out* of control. This further highlights the interaction between awareness, the sense of agency and the sense of self.

#### ***Interactions with Technology.***

An important challenge in the development of human-computer interfaces lies precisely in maximising the agent’s experienced control over the computer (Limerick, Coyle, & Moore, 2014). Yet, there may be a simultaneous need to reduce explicit awareness of the interface system and of the self, to produce an immersive experience, such as in virtual reality settings, in turn blurring the boundaries between self and interface/computer. Interestingly, interactions with technology raise another issue for SoA – the effects of task automation by computers or machines, e.g. in aviation, or factories. While this may be helpful for completing the task, the reduced engagement of the user, and increased uncertainty over what the user can control, can lead to a reduced SoA (Berberian, Sarrazin, Le Blaye, & Haggard, 2012). This has important implications for the behaviour of users, as well as for establishing personal and legal responsibility.



### ***Responsibility and the Law.***

The belief that people can control their behaviour is not only relevant to individuals, but is also at the heart of our societal notions of personal and moral responsibility, and thus our legal system (Moretto, Walsh, & Haggard, 2011; Spence, 2009). Clarifying the mechanisms underlying SoA can thus potentially inform these concepts. Most legal systems are primarily concerned with establishing the facts of agency, that is, determining *who* caused a certain event, or committed a crime. Nonetheless, the subjective *experience* of agency is also considered important. To assign guilt, most legal systems take into account: whether there was malicious intent; whether the agent could control their actions (vs. coercion, for example); and whether they were able to understand the consequences of action. This may result in different charges, such as manslaughter for an accidental death, instead of murder. Or different defences may be applicable, namely the insanity defence if the agent was in a psychotic episode, which affected their understanding of their action consequences, or led them to believe they were being controlled by external forces.

### ***Psychiatric and Neurological Disorders.***

In fact, a number of clinical disorders have been associated with disturbances in SoA. Most typically associated is schizophrenia, as patients with delusions of control believe that their thoughts and actions are controlled by external forces (Frith, 2005; Hauser et al., 2011; Moore & Fletcher, 2012). Moreover, depression can be associated with a loss in the sense of control over one's life and behaviour (Bandura, 1991); while feelings of incompleteness in the performance of actions have been found in obsessive-compulsive disorder (OCD, Belayachi & Van der Linden, 2010; Gentsch, Schütz-Bosbach, Endrass, & Kathmann, 2012). Finally, neurological lesions can also result in disturbances in SoA, such as anarchic hand syndrome, which can follow to damage to supplementary motor area (SMA), and in which patients feel that they cannot control the movements of the hand contralateral to the lesion (Marchetti & Della Sala, 1998). Therefore, research on SoA has potential implications for the development of new therapeutic targets or interventions.

### 1.1.2. Summary

The significance and pervasiveness of the experience that people can control their actions and are responsible for their consequences is indisputable. Nevertheless, the mechanisms that underlie its emergence remain unclear. How do our intentions, actions, and their consequences come to be linked in this apparent voluntary causal chain? Moreover, we might wonder what parts of this chain are related to SoA. Does agency depend on an inference triggered by the perception of the outcome to retrospectively link it to our action? Or does agency also rely on internal signals associated with the intention itself, such as a metacognitive perception of action selection, to prospectively link it with the action?

Much research has investigated the retrospective link, and shown that SoA is based on a comparison between predictions, or expectations, about action outcomes and observed outcomes. If these match, SoA is high, but a mismatch leads to a loss of agency. While this comparison can involve inferential processes (Wegner & Wheatley, 1999), it can also rely on internal, sensorimotor predictive processes (Blakemore, Wolpert, & Frith, 2002). Since the outcome must be known for the comparison to take place, both of these mechanisms are essentially *retrospective*.

Thus far, literature on SoA had largely neglected the link between intention and action, that is, the role of decision making, or action selection, processes. Recent studies have indeed shown that the fluent, or easy, action selection is associated with a stronger SoA, than dysfluent, or difficult, selection (Chambon, Sidarus, et al., 2014). This reveals that a metacognitive signal related to the fluency of action selection processes can influence SoA. As these processes clearly precede the outcome, and even the action itself, they offer a unique window to investigate a *prospective* component of SoA.

### 1.1.3. Motivation for this research

The present research aimed to explore further the prospective aspect of SoA, by investigating the role of action selection processes. We focused on situations in

which agency is especially relevant and informative for guiding current and future behaviour, and which have been largely neglected in the previous literature. For example, when interacting with new environments, the instrumental relation between actions and outcomes must be learned. When little is known about action outcomes, other cues, namely prospective ones, may be used to support our SoA and guide instrumental learning.

Moreover, we often face situations in which there are suggestions for alternative actions, or we are unsure about how to make a decision. Metacognitive monitoring of action selection can recruit cognitive control processes in order to resolve such conflict or uncertainty, and adapt future behaviour. Such signals may additionally inform the SoA. Finally, social contexts can increase ambiguity about the authorship of outcomes (Frith, 2014; Wegner & Wheatley, 1999). However, the presence of alternative agents may also challenge the SoA by increasing the complexity and uncertainty of decision making processes, as one needs to consider the potential actions of others, the outcomes of those actions, and so on. Therefore, prospective aspects of SoA may also be implicated in social situations.

The following literature review will start by considering the ways in which we can measure SoA, as this has implications for our models of SoA. Such theoretical models of SoA, and the cues that inform it, will then be reviewed. Next, we will consider how the SoA may be embedded in psychology in general, and its functional interaction with other aspects of the mind. We will then highlight the gaps and limitations of the present literature on SoA. Finally, we will demonstrate how the present research aimed to enhance our understanding of the cognitive and neural mechanisms of a prospective SoA.

## 1.2. Research on Agency

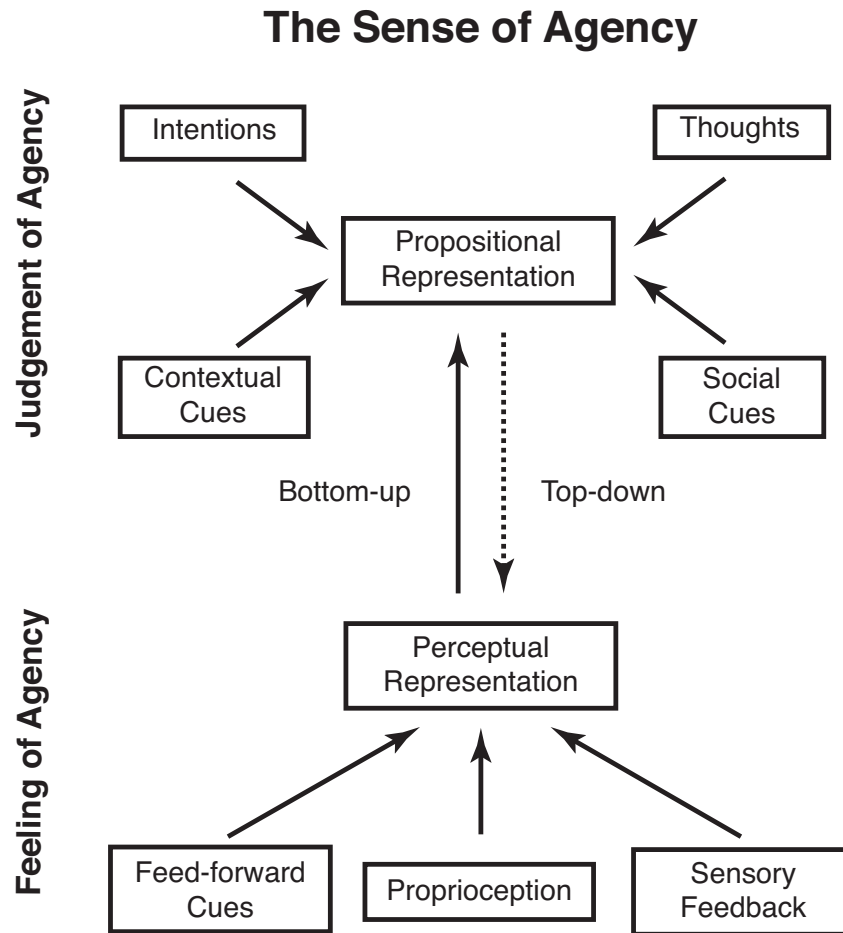
### 1.2.1. How can we measure the sense of agency?

A variety of measures have been used to investigate the subjective experience of agency. Before considering them in detail, some conceptual distinctions are needed. First, there are two, interacting, levels of the SoA (Synofzik et al., 2008; see **Figure 1.1**). The feeling of agency refers to a pre-reflective, non-conceptual representation, related to internal sensorimotor signals. The judgement of agency is a reflexive, conceptual representation, which draws from the feeling level, but also encompasses other cognitive cues such as beliefs and contextual information, and inferential processes. This highlights the difficulties found in measuring the SoA, as judgments of agency can be assessed through explicit reports, but targeting the feeling of agency requires more implicit measures. Moreover, we can distinguish between an *attributional* aspect of SoA, linked to determining the authorship of outcomes (“I did that”), and an *instrumental* aspect of SoA, concerned with the relation between specific actions and specific outcomes (“I did *that*”; Chambon, Filevich, & Haggard, 2014).

#### ***Explicit measures.***

Many studies have asked people directly to report on their experience, i.e. probing judgements of agency. Some have focused on the authorship of action outcomes, asking participants to assess how much they felt an outcome was caused by their action, vs. by another agent, either with dichotomous responses (Bednark & Franz, 2014; Farrer, Frey, et al., 2008; Ritterband-Rosenbaum, Karabanov, Christensen, & Nielsen, 2014), or with continuous scales (Aarts, Custers, & Wegner, 2005; Damen, van Baaren, & Dijksterhuis, 2014; Kühn et al., 2011; Sato & Yasuda, 2005; Spengler, von Cramon, & Brass, 2009). Yet, even when we know outcomes are caused by our actions, we may feel more or less control over them. Thus, other studies have asked participants how much control they felt over their action, or action outcome (Chambon & Haggard, 2012; Dewey, Seiffert, & Carr, 2010; Linser & Goschke, 2007; Metcalfe & Greene, 2007; Nahab et al., 2011; Sebanz & Lackner, 2007). Explicit reports are crucial for developing an account of SoA that is in line with our conscious

subjective experience. However, they are also known to be susceptible to cognitive biases, such as social desirability, and the questions used to elicit the reports are subject to interpretation from participants.



**Figure 1.1.** Two levels of the sense of agency. Sense of agency receives many inputs, but different cues are integrated at a non-conceptual level – the feeling of agency, and at a conceptual level – judgement of agency. The feeling level forms the basis of the judgement level (bottom-up effect), but the judgement level can also influence the feeling level, based on beliefs or prior knowledge (top-down effect). Figure adapted from Synofzik et al. (2008).

### ***Implicit measures.***

Alternatively, implicit measures, linked to sensorimotor phenomena, have been proposed to offer a marker of the feeling of agency. Voluntary actions can lead to a sensory attenuation of outcomes (Blakemore, Wolpert, & Frith, 1998), linked to motor prediction (see Hughes, Desantis, & Waszak, 2012 for a review). Relative to externally-triggered stimuli, self-produced stimuli are associated with an attenuation

in perceived intensity (e.g. Desantis, Weiss, Schutz-Bosbach, & Waszak, 2012), as well as in neural responses, as seen in fMRI (Blakemore et al., 1998), or in EEG (Gentsch & Schütz-Bosbach, 2011; Hughes, Desantis, & Waszak, 2013). Another implicit measure of agency is linked to the perceptual compression of the temporal interval between action and outcome, termed intentional binding (Haggard, Clark, & Kalogeras, 2002). Intentional binding was stronger when participants chose freely when to act, relative to baseline, and to TMS induced movements (see Moore & Obhi, 2012 for a review). Finally, cortico-spinal excitability has also recently been suggested to index SoA, as it is enhanced for visual feedback matching intentional actions (Weiss, Tsakiris, Haggard, & Schütz-Bosbach, 2013).

### ***Relation between explicit and implicit measures.***

Interestingly, explicit and implicit measures of SoA can sometimes be dissociated (Ebert & Wegner, 2010; Moore, Middleton, Haggard, & Fletcher, 2012). They can also be dissociated when comparing different tasks to probe individual differences in SoA (Dewey & Knoblich, 2014; Saito, Takahata, Murai, & Takahashi, 2015), even when comparing different implicit measures (Dewey & Knoblich, 2014). As mentioned before, the SoA is complex and multifaceted, thus different measures may reflect different aspects of the experience of agency. Combining them in the same study may offer new insights into the mechanisms that underlie SoA.

## **1.2.2. How does the sense of agency come about?**

### **1.2.2.1. Linking Intentions and Outcomes**

#### ***Instrumental learning.***

In the early 20th century, the notion of operant behaviour, was established by Thorndike in his Law of Effect (Thorndike & Bruce, 1911). This referred to the ability of animals to learn associations between certain actions and outcomes. Positive outcomes reinforced a link with the respective causal behaviour, whereas negative outcomes weakened such associations. Although often neglected in the agency literature, this is the core ability that underlies our agentic capacities. This

instrumental learning mechanism allows animals and humans to drive their behaviour on the basis of desirable goals, and intervene in the environment autonomously, i.e. not only in reaction to a stimulus (de Wit & Dickinson, 2009). Hence, on this widely held view, compatible with our common sense experience of agency, goals or intentions initiate the appropriate actions in order to obtain desired outcomes.

***Ideomotor theory.***

Similarly, William James' ideomotor theory (1890) proposed that the representation of a desired outcome was a necessary precondition for voluntary actions, and that, in fact, it was the anticipatory representation of the outcome that drove action selection and execution. Recent versions of this proposal emphasise that action and perception share representational resources (Elsner & Hommel, 2001; Schütz-Bosbach & Prinz, 2007). In this view, as we learn about contingent actions and outcomes, their association becomes bidirectional, such that activation of one enhances activation of the other. This proposal can account well for known interactions between motor and perceptual systems. A limited capacity in the system can impair the co-activation of action- and perceptually-related representations and cause a reduction in task performance (Musseler & Wuhr, 2002). Conversely, studies have shown that the two representations can be assimilated and actually boost task performance when there are small temporal asynchronies between the two. An example is given in response priming induced by stimuli that were previously learned as effects (Kiefer, Sim, Helbig, & Graf, 2011; Nikolaev, Ziessler, Dimova, & van Leeuwen, 2008; Schubö, Prinz, & Aschersleben, 2004; though see Grosjean, Zwicker, & Prinz, 2009). Thus, this theory provides another account of how intentions become associated with their outcomes.

***An inferential account of sense of agency.***

Another more recent view that stresses goal representations stands, however, in stark contrast with these previous accounts by proposing that the experience of agency is really an illusion (Wegner, 2004). In his theory of "apparent mental causation", Wegner (2004) argues that, rather than relying on privileged access to

internal action-related signals, SoA results from a retrospective inference about the causal relationship between anticipatory thoughts and action outcomes. This is guided by 3 principles: the thought should precede the outcome (“priority”) and be consistent with it (“consistency”), and there should be no other obvious cause for the outcome (“exclusivity”). This theory is based on a number of studies demonstrating that people can hold erroneous beliefs about their agency over certain events. For example, in the “I Spy” experiment, participants moved a pointer around a screen containing various images together with a confederate and stopped occasionally. Participants were also incidentally exposed to words that could refer to objects on the screen. Results showed that if there was consistency between a word they had recently heard and the image at which the pointer stopped, participants could self-attribute the stopping event, even though a confederate had caused it (Wegner & Wheatley, 1999). Therefore, a match between preceding thoughts and outcomes leads to self-agency attributions; while mismatches lead to a reduced sense of authorship over outcomes. These findings are important for theories of SoA as they highlight that we can make errors in determining our agency. However, this account only addresses a high-level cognitive process of attribution of action consequences to particular actors.

The accounts described above emphasise the role of the external consequences of action. Internal signals associated with selecting and executing the action are neglected. In other words, these accounts focus on the representation of the end, or goal, states, while downplaying the representation of the means by which the ends are achieved. Therefore, these accounts imply that SoA depends on the monitoring of external outcomes, rather than on monitoring internal signals. Moreover, recent frameworks of SoA emphasise that it can involve a graded sense of controlling one’s actions and their outcomes, in addition to self vs. other causal attributions (Chambon, Sidarus, et al., 2014; Farrer, Valentin, & Hupé, 2013; Pacherie, 2007; Synofzik et al., 2008). This sense of control depends on internal signals related to action, thus reclaiming the role of the means through which we achieve desirable ends as a constitutive part of SoA.

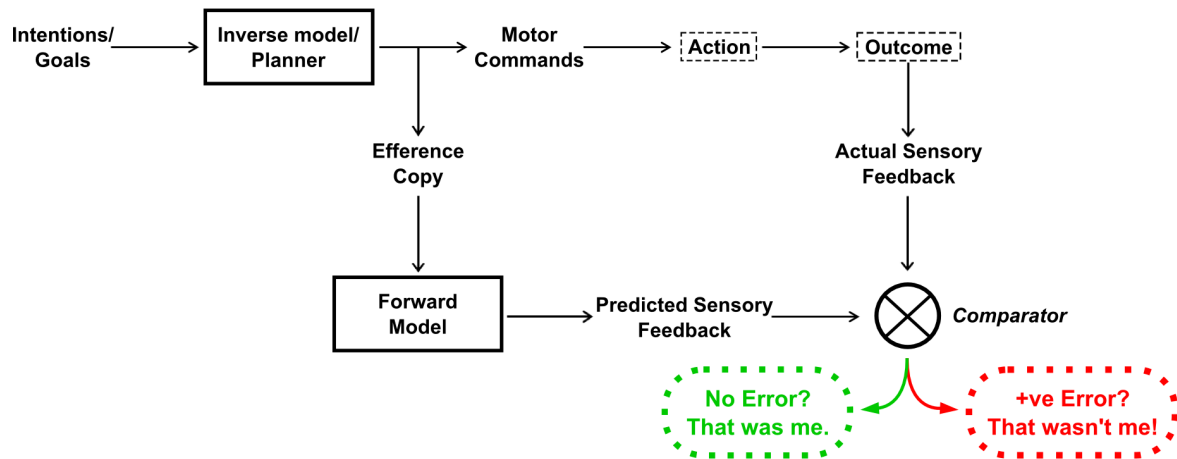


### 1.2.2.2. Linking Actions and Outcomes

#### ***The comparator model of sense of agency.***

In fact, the contribution of implicit and embodied signals involved in action control and monitoring to the SoA have been well established (Blakemore, Frith, & Wolpert, 1999; Sato, 2009; Tsakiris, Haggard, Franck, Mainy, & Sirigu, 2005); for a review see Haggard & Tsakiris, 2009). Computational models of sensorimotor control have provided a useful framework for understanding how sensorimotor signals might be relevant to agency. A model by Wolpert and colleagues proposes the following (Wolpert, Ghahramani, & Jordan, 1995; see Figure 1.2 below for the schematic of an extended model). Intentions or goals are transformed into motor commands in planners or *inverse* models. The motor commands are then sent to muscles in the body, but an efference copy, i.e. a copy of the motor command, is passed to a *forward* model, which predicts the sensory consequences of the movement. In motor control, this predictive mechanism is critical for rapid movement corrections as it can bypass the time delay of sensory feedback. Additionally, sensory predictions are passed to a *comparator*, which also receives actual sensory feedback. This subtracts the prediction from sensory feedback. Therefore, activity in sensory systems is proportional to prediction errors. This prevents sensory systems from becoming overloaded with uninformative sensory input. Moreover, this results in a sensory attenuation for predicted sensory stimulation.

This model has subsequently been extended beyond action control to account for how we can distinguish self-generated from externally-triggered sensory stimulation (Blakemore et al., 1999; see Figure 1.2). According to the model above, self-triggered sensory stimulation will be attenuated because we can predict it based on our action, whereas externally-triggered stimulation cannot be predicted as accurately. Hence, the absence of an error signal produced by the comparator implies that I caused the sensory stimulation; whereas prediction errors, due to mismatching sensory input and predictions, may indicate that I did not cause that sensory event.



**Figure 1.2.** Schematic of the computational model for sensorimotor control. Adapted from (Blakemore et al., 1999). The output of the comparator can be used to determine the cause of sensory input. If prediction and feedback match, there is no error signal, and I can self-attribute the cause of the sensory stimulation. But if prediction and feedback mismatch, a prediction error signal will indicate that I was not the cause of the sensory input.

In a seminal study, Blakemore and colleagues (Blakemore, Wolpert, & Frith, 1998) investigated why people cannot tickle themselves, by contrasting the neural activation for self- and externally-produced tactile stimulation. They found that activity in the somatosensory cortex was reduced for self-produced stimulation. Moreover, they reported that increasing the temporal asynchrony between the subject making a left-hand movement and receiving tactile stimulation on the right hand also increased subjective ratings of “ticklishness”. Many studies have provided support to the role of sensory attenuation in self-other distinctions. Similar effects have been shown in the auditory domain (Baess, Widmann, Roye, Schröger, & Jacobsen, 2009; Kühn et al., 2011; Weiss, Herwig, & Schütz-Bosbach, 2011), as well as for visual stimuli (Gentsch, Kathmann, & Schütz-Bosbach, 2012; Gentsch & Schütz-Bosbach, 2011; Hughes & Waszak, 2011, see Hughes, Desantis, & Waszak, 2012 for a review). Therefore, sensory attenuation has been proposed to reflect an implicit, pre-reflective feeling of agency (Blakemore et al., 2002; Gentsch & Schütz-Bosbach, 2011; Sato, 2009; Synofzik et al., 2008).

Another relevant aspect of this sensorimotor model to SoA is that the output of the comparator could be used not only to detect but also measure the discrepancy between predictions and outcomes. Therefore, the output of the comparator can serve to index the degree of control experienced over action outcomes, and not only for making self-other distinctions (see Synofzik, Vosgerau, & Newen, 2008 for a review). Consistently, increases in the discrepancy between action and visual feedback are associated with a reduction in perceived control (Dewey et al., 2010; Metcalfe & Greene, 2007; Nahab et al., 2011).

### ***A hierarchical framework of intention specification and monitoring.***

This “comparator model” of SoA has proved useful beyond low-level sensorimotor control, as intentions, action, and outcomes can be conceived at different levels. Pacherie (2008) proposes three levels of intention can be distinguished in the process of action specification. Distal intentions regard distal goals, such as turning on the lights in a dark room. Proximal intentions regard the more proximal means available to satisfy the distal goals, like flipping a light switch. Finally, motor intentions consist of the motor commands necessary to enact the proximal intentions. The author further proposes that these different, hierarchically organised, levels of intentions are associated with different levels of monitoring and control processes, and thus multiple levels of predictions and observed outcomes to compare.

### ***Metacognitive contributions to sense of agency.***

Importantly, monitoring at these different levels of intentions and outcomes can have different influences to SoA. Using a videogame-like computer task, Metcalfe and colleagues (Metcalfe, Eich, & Miele, 2013) manipulated proximal and distal outcomes of action independently. The task required catching falling Xs while avoiding Os, by moving a box along a horizontal bar, with the computer mouse. When an item was caught, it disappeared, and auditory feedback indicated the accuracy of the catch. Adding a discrepancy between the movements of the mouse and the movements of the box on the screen interfered with proximal outcomes (i.e. the action at the distal level). Distal outcomes were manipulated by reducing the contingency between

touching the item with the box and the item disappearing, from 100% to 75%. Results showed that disruptions at the proximal level of action led to a large reduction in judgements of agency, whereas only a small reduction was observed when the distal outcomes were not consistently achieved. Moreover, results showed that judgements of agency were also positively related to the perceived overall performance in the game, in terms of how many Xs were caught (cf. Metcalfe & Greene, 2007). The authors argued that the SoA is metacognitive, as it involves monitoring internal signals associated with sensorimotor and cognitive processes, rather than only relying on the perception of external events (see also section 1.3.1 below). Importantly, these findings further emphasise the importance of the means through which our goals are achieved (see also Caspar, Cleeremans, & Haggard, 2015).

The evidence reviewed in this section clarifies the link between action and outcome, and reinforces the importance of internal sensorimotor signals in the experience of agency. This research has shown that we have (some) direct access to internal bodily signals, and do not *only* rely on conceptual, inferential processes to assess our agency. However, this does not preclude the influence of *any* inferential processes to SoA. Nevertheless, these accounts of agency are still dependent on a representation of the outcome. Only when the outcome is known can it be compared with a prediction. Therefore, these processes can only inform SoA *retrospectively*. Moreover, the accounts discussed so far fail to recognise the importance of decision making, or action selection, processes to SoA. For example, if an archer releases her arrow accidentally, e.g. due to a loud noise, and hits the bulls eye, she will likely not feel in control of that outcome, even though it matches her goal, and she knows she caused it. Consequently, we can hypothesise that there is a prospective component of agency associated to the process of linking our intentions to our actions.

### 1.2.2.3. Linking Intentions to Actions

#### ***Prospective contributions to sense of agency.***

Prospective cues to agency do arise even before action execution. Recent studies have found that action selection processes can influence the SoA (Chambon & Haggard, 2012; Chambon, Moore, & Haggard, 2014; Chambon, Wenke, Fleming, Prinz, & Haggard, 2013; Sebanz & Lackner, 2007; Sidarus, Chambon, & Haggard, 2013; Wenke, Fleming, & Haggard, 2010; Wenke, Gaschler, Nattkemper, & Frensch, 2009, see (Chambon, Sidarus, et al., 2014). For example, subliminal priming can be used to manipulate action selection (Eimer & Schlaghecken, 2003; Lingnau & Vorberg, 2005; Vorberg, Mattler, Heinecke, Schmidt, & Schwarzbach, 2003). When primes are congruent with targets, action selection is easy, but when primes are incongruent with targets, action selection is impaired, as evidenced by slower reaction times (RTs) and more errors. In one study, participants responded to directional arrows, which were preceded by subliminal primes (Wenke et al., 2010). Actions triggered the appearance of coloured circles, after a variable delay. Participants were instructed to attend to the relation between actions and outcomes. Results showed that participants gave lower judgements of agency for outcomes that followed incongruently primed actions, compared to congruently primed actions. This shows that dysfluency, or difficulty, in selecting a correct response leads to a reduction in SoA.

Importantly, this effect is independent of outcome monitoring, as outcomes were equally predictable across priming conditions. Moreover, fluency effects cannot be explained by RT monitoring (Chambon & Haggard, 2012; Chambon et al., 2013; Sidarus et al., 2013). In subliminal priming, a short interval between prime and target induces the effects previously described, termed “positive compatibility effect” (PCE). However, a longer prime-target interval can lead to a “negative compatibility effect” (NCE), wherein congruent priming leads to slower RTs and more errors than incongruent priming (Eimer & Schlaghecken, 2003; Lingnau & Vorberg, 2005). Taking advantage of this divergence of effects, Chambon and Haggard (2012) showed that under both PCE and NCE conditions, incongruent priming led to a

reduction in SoA, compared to congruent priming. The influence of action selection on SoA is not linked to RT monitoring.

These studies show that a metacognitive signal about the fluency of action selection processes can inform SoA *prospectively*. It has been suggested that the fluency of action selection may serve as an advance predictor of successful action (Haggard & Chambon, 2012). We may learn through experience that easy and well performed actions are more likely to yield the desired, or expected, outcomes than difficult or disrupted actions. Indeed, expertise is associated with better outcome prediction, e.g. in expert baseball players (Gray, Beilock, & Carr, 2007). We may learn to use a metacognitive feeling (Arango-Muñoz, 2010) of fluency as a heuristic device for estimating our SoA before the outcome is known.

#### 1.2.2.4. Integrating Multiple Cues

As reviewed above, a number of processes have been found to influence the SoA. Recent frameworks highlight the integrative nature of SoA, which is informed by a variety of cues (Gallagher, 2012; Haggard & Tsakiris, 2009; Moore & Fletcher, 2012; Synofzik et al., 2008). In addition to inferences, beliefs and contextual information, internal signals are a critical component of SoA (see **Figure 1.1** above). Moreover, different internal signals can arise at different times in the voluntary action chain (Chambon et al., 2013; Farrer et al., 2013).

Moore and Fletcher (2012) have proposed that Bayesian cue-integration models may help in understanding this integrative process. Here, cues would be weighted differently according to their reliability and availability. These models have proved useful in understanding the integration between sensorimotor predictions and inferential processes (Wolpe, Haggard, Siebner, & Rowe, 2013). However, the reliability of cues may change over time, for example during instrumental learning. For example, under conditions of ambiguity in agency attribution, it has been shown that self-agency reports were associated with recent self-agency experiences (Bednark & Franz, 2014). Moreover, these proposals still cannot account for the

integration of cues over time (Farrer et al., 2013), nor for how the specific value of a cue may alter its weighing function (Sidarus et al., 2013).

#### **1.2.2.5. Neural Substrates of Sense of Agency**

Finally, it is worth briefly considering the neural substrates of our experience of agency (see David, Newen, & Vogeley, 2008 for a review). Many studies have implicated the regions in the parietal cortex in detecting mismatches between predicted and observed sensory feedback (Fink et al., 1999; Miele, Wager, Mitchell, & Metcalfe, 2011; Spengler et al., 2009, 2009; Sperduti, Delaveau, Fossati, & Nadel, 2011; Yomogida et al., 2010), namely the angular gyrus (AG; Chambon, Moore, et al., 2014; Farrer et al., 2003; Farrer, Frey, et al., 2008; Farrer & Frith, 2002; Khalighinejad & Haggard, 2015; Nahab et al., 2011; Spengler et al., 2009; Tsakiris, Longo, & Haggard, 2010). Interestingly, increased activity in the AG was related to the conscious experience of a loss of agency, in association both with prospective, i.e. dysfluent action selection (Chambon et al., 2013), and retrospective cues to SoA, i.e. prediction-outcome mismatch (Farrer, Frey, et al., 2008). This suggests the AG may be involved in the online monitoring of agency. These findings support a view of self-agency as a default, with disruptions in the smooth flow from intentions, to actions, to outcomes, resulting in a loss of agency.

Additionally, regions associated with initiating and monitoring voluntary action have been implicated. Disruption of the pre-supplementary motor area (pre-SMA) has been linked to a reduction of SoA (Cavazzana, Penolazzi, Begliomini, & Bisiacchi, 2015; Javadi, 2015; Moore, Ruge, Wenke, Rothwell, & Haggard, 2010). This suggests a role for the pre-SMA in supporting a strong SoA (Tsakiris et al., 2010), similarly to the SMA (Kühn, Brass, & Haggard, 2012; Yomogida et al., 2010), and the dorsolateral prefrontal cortex (dlPFC; Chambon et al., 2013; Khalighinejad, Di Costa, & Haggard, 2016).

Finally, other frontal regions have been implicated in the more reflexive aspects of SoA. The anterior PFC has been associated with stronger judgements of agency (Miele et al., 2011), and was previously implicated in metacognition of action

(Fleming & Dolan, 2012; Fleming, Weil, Nagy, Dolan, & Rees, 2010), as well as self-knowledge and theory of mind, i.e. attributing mental states to others, more generally (Amodio & Frith, 2006). On other hand, the dorso-medial frontal cortex has been linked to external agency attributions (Spengler et al., 2009; Sperduti et al., 2011), consistent with its involvement in mentalizing tasks, i.e. attributing mental states to others (Amodio & Frith, 2006).

### **1.3. Sense of Agency and Other Mental Functions**

In the above literature review on SoA we saw that much has been learned about how we link actions to outcomes, and self-attribute agency over those outcomes. However, the role of executive functions such as planning, decision making and cognitive control have been largely neglected, even though these are typically considered critical for the exercise of our agentic capacities (Schiffer, Waszak, & Yeung, 2015). The field of metacognition, in particular in relation to decision making, has been concerned with understanding how we monitor our cognitive processes in order to flexibly adapt our cognition and behaviour (Yeung & Summerfield, 2012). Recent work on prospective SoA highlights the need to bridge these two fields for a more complete understanding of human agency. Furthermore, the literature discussed above has also largely neglected how social context may influence SoA, and behaviour. The present section will briefly review these two topics, which are implicated in the research conducted for this thesis.

#### **1.3.1. Metacognition**

Metacognition can be simply defined as involving a meta-level cognitive process which is *about* an object-level cognitive process (Nelson & Narrens, 1990). It has been suggested that there are two interrelated levels of metacognition: a) a lower-level involved in the monitoring and control of cognitive processes; and b) a higher-level meta-representational level that interprets behaviour based on beliefs and theories (Arango-Muñoz, 2010; Koriat, 2000). “Epistemic feelings” may arise from monitoring processes and can be used to adjust behaviour online (Proust, 2008).



However, while both levels may be associated with explicit, conscious metacognition, monitoring and control processes may remain implicit (Fleming & Dolan, 2014).

#### **1.3.1.1. Conflict Monitoring**

Monitoring action selection allows the recruitment of cognitive control processes when needed, in order to adjust subsequent behaviour. For example, when a target stimulus appears surrounded by incongruent distractors, which activate a competing response alternatives, this response conflict leads to a slowing of RTs (e.g. Eriksen & Eriksen, 1974). Yet, this congruency effect is reduced in trials following response conflict, relative to following easy, congruent trials (Gratton, Coles, & Donchin, 1992; see Egner, 2007 for an overview). That is, detection of response conflict triggered behavioural adaptation, in order to improve performance. Although debate is ongoing, some studies have shown that conflict adaptation can also occur for unconsciously triggered response conflict (Atas, Desender, Gevers, & Cleeremans, 2015; van Gaal, Lamme, & Ridderinkhof, 2010, but see Desender, Van Lierde, & Van den Bussche, 2013). Interestingly, a recent study, using unconsciously triggered conflict, suggested that a conscious *experience* of conflict, or difficulty in action selection, was necessary for conflict adaptation (Desender, Opstal, & Bussche, 2014).

#### ***Neural markers of conflict.***

At a neural level, the anterior cingulate cortex (ACC) is thought to be involved in conflict monitoring, and to trigger cognitive control functions associated with more frontal regions, such as the dlPFC (Botvinick & Braver, 2015; Holroyd & Yeung, 2012). Moreover, ACC-mediated conflict monitoring can also be identified in event related potentials (ERPs; for a review, see Larson, Clayson, & Clawson, 2014). In ERPs locked to target onset, response conflict leads to a large N2 component, a negative potentials peaking around 250-300ms after the target, at fronto-central sites (Kopp, Rist, & Mattler, 1996). This N2 component is thought to index conflict detection and resolution (Larson et al., 2014).

**Error monitoring.**

Conflict monitoring theory has also been used to explain error monitoring and detection (Yeung, Botvinick, & Cohen, 2004). As accumulation of evidence for the correct response may continue after selecting a given action, post-decisional processing can reveal that the chosen response was incorrect, resulting in a conflict between the correct and the executed response (Yeung & Summerfield, 2012). In ERPs locked to the action, an error-related negativity (ERN) component emerges immediately after error commission (0-100ms), in comparison to correct responses (see Larson et al., 2014 for a review). In correct trials, a correct-related negativity (CRN) has been associated with task difficulty/uncertainty (Pailing & Segalowitz, 2004; Scheffers & Coles, 2000). Errors and the ERN have also been associated with ACC activity (Carter et al., 1998; Charles, Van Opstal, Marti, & Dehaene, 2013). Therefore, it has been suggested that the target-locked N2 and the action-locked ERN components reflect pre- and post-decisional conflict monitoring, linked to the ACC (Larson et al., 2014; Yeung & Summerfield, 2012).

**1.3.1.2. Confidence**

In contrast to research on error monitoring, models of confidence judgements have typically emphasised a role for pre-decisional processing (Yeung & Summerfield, 2012). Confidence is related to the speed of accumulation of evidence, as well as to the balance of evidence between response alternatives at the time of action (Kiani, Corthell, & Shadlen, 2014). However, post-decisional processing also influences confidence (Boldt & Yeung, 2015; Resulaj, Kiani, Wolpert, & Shadlen, 2009; Scheffers & Coles, 2000), showing that confidence judgements and error monitoring can both draw on the same post-decisional processing. For example, a study (Boldt & Yeung, 2015) varied the difficulty of a perceptual discrimination task, and obtained judgements on a six-point scale, ranging from certainly correct, to maybe correct, to certainly wrong. Results showed that increasing confidence in having made an error was associated with more negative CRN/ERN amplitude, as well a more positive Pe amplitude (another error-related ERP component found around 300ms after action).

Moreover, models of confidence judgements have often neglected the role of action and the motor system, where evidence accumulation can occur in parallel to perceptual processing (Cisek, 2007). A recent study showed that TMS stimulation of the premotor cortex associated with the unchosen response disrupted metacognitive accuracy in perceptual confidence judgements (Fleming et al., 2014). Moreover, these effects were found for stimulation both before and after the action. Therefore, confidence judgements rely on late-stage metacognitive processes, which are influenced by action-specific signals. These findings are particularly relevant to the present thesis, as they are consistent with the aforementioned influence of fluency in action selection to the sense of agency, which has also been linked with post-decisional processing (Chambon, Moore, et al., 2014).

#### **1.3.1.3. Fluency**

Finally, the research considered above is also relevant to the widely studied effects of fluency on a variety of judgements, such as confidence, liking or familiarity (Alter & Oppenheimer, 2009). These effects have been associated with fluency at many levels of processing, such as perceptual, cognitive, linguistic, or memory-based. The experience, of feeling, of fluency is thought to be based on a qualitative signal about information processing, and can be broadly defined as a continuum, ranging from fluent or effortless, to dysfluent or effortful processing. Note that response conflict can lead to a conscious experience of *dysfluency* (Desender et al., 2014; Morsella et al., 2009). These experiences, or feelings are often vague, especially about their sources (Winkielman, Ziembowicz, & Nowak, 2015). Therefore, fluency/dysfluency experiences can “leak” into judgements, even if they may not be a relevant cue. It has also been argued that conflict (or dysfluency) may serve as an aversive signal (Botvinick, 2007). Consistently response conflict can lead to more negative affective judgements of subsequent neutral stimuli (Fritz & Dreisbach, 2013). Importantly, when fluency experiences can be attributed to an irrelevant source, they no longer influence judgements (cf. Alter & Oppenheimer, 2009). Therefore, while fluency may often be a useful heuristic cue (Whittlesea & Leboe, 2003), it will not be taken into account if considered uninformative.

The work on metacognition reviewed in this section offers some insight into how action selection fluency influences the SoA. The work on conflict monitoring suggests that neural signals associated with response conflict detection and resolution could form the basis of the cue to SoA. Relatedly, the work on confidence suggests that similar, post-decisional, metacognitive processes may inform both confidence and agency judgements. Moreover, fluency effects on SoA could be linked to a general heuristic that would enhance SoA, or due to the affective consequences of response conflict.

### **1.3.2. Social Aspects of Agency**

The experience of agency can be especially relevant in social contexts, as the presence of other agents can increase ambiguity in agency attribution. Moreover, concepts of personal and moral responsibility become more salient. These contextual effects may thus have important consequences to our SoA, as well as to our behaviour. In fact, the influence of social context on behaviour is well documented in the social psychology literature. For example, in emergency scenarios, the likelihood of someone helping decreases with the number of bystanders (Darley & Latane, 1968); when working in a group, people put in less effort than if they were working alone (Karau & Williams, 1993). These effects are thought to result from a diffusion of responsibility, in which individuals feel the responsibility for action lies with others (Bandura, 1991). Yet, this could merely reflect a post-hoc justification, related to self-serving biases, such as maintaining a positive self-image, rather than involving online changes in SoA.

Interestingly, a recent study has shown that being coerced into giving electric shocks to others leads to a reduction in intentional binding, and an attenuation of outcome processing, relative to a free choice condition (Caspar, Christensen, Cleeremans, & Haggard, 2016). This shows that social context can influence low-level, implicit aspects of SoA, in a condition in which social desirability and self-serving biases were thought to underlie a reported reduction in responsibility. Under coercion, responsibility is displaced onto another agent (Bandura, 1991), thus agency

attribution may become more ambiguous. This is clearly also the case in a group work scenario. However, the help of one person may suffice in emergency situations, therefore, changes in decision making processes may be more relevant than attribution ambiguity. Social situations are associated with more complex decision making, and uncertainty, since the potential behaviour of other agents needs to be considered. Consequently, in line with the fluency effects mentioned above, it could be hypothesised that an increased difficulty in decision making could result in a reduction in SoA, and in the motivation to act.

## **1.4. Limitations of the Previous Literature**

### **1.4.1. Focus on retrospective agency attribution**

The literature on SoA reviewed above shows that the role of action selection processes has been largely neglected. In particular, most of the literature that directly discusses “agency” involves explicit retrospective judgements to attribute agency to the self or others, often creating ambiguity through social situations. This is the easiest aspect of agency to measure, but it is only one of several varieties of agency experience. Our subjective experience of agency shows us that we can feel more or less in control of our actions and their outcomes, even when we know we have caused them. Moreover, many studies on SoA fail to adequately distinguish between different levels of goal and action representations. Thus, the apparent action in some studies (e.g. moving a cursor on a screen to reach a target) may be considered an outcome in others. A hierarchical organisation of intentions, actions and outcomes, linked to successive levels of action specification (e.g. Pacherie, 2008), may help gain a better understanding of the interaction between different levels, and cues to agency. Finally, previous models imply that SoA only can only arise once the outcome is known. This ignores the possibility that SoA involves an online monitoring of multiple cues, which can become available at different times in voluntary action, and the subjective experience that SoA begins already during action selection. In general, this review of the literature shows that our understanding of the sense of agency has been largely like looking for one’s keys under the

streetlamp, irrespective of where one lost them. The concept of agency has become restrictively defined by the most intuitively obvious method of measuring it, namely explicit agency attribution. As a result, other varieties of agency experience, and other contributors to sense of agency, have been under-emphasised.

#### **1.4.2. Current prospective models remain unclear**

Recent work has revealed a prospective contribution to SoA, linked to action selection fluency. However, the concept of *fluency* remains unclear, especially since it has been dissociated from monitoring RTs (Chambon & Haggard, 2012). It also remains unclear which stages, or aspects, of action selection could contribute to SoA. In addition to pre-motor response conflict, could processing fluency, or uncertainty also influence SoA? Additionally, the integration between prospective, action-related, and retrospective, outcome-related, cues remains poorly understood (but see Sidarus et al., 2013). Finally, can such fluency effects on SoA be found if conscious stimuli are used to manipulate action selection?

#### **1.4.3. Does agency reflect fluency or effort?**

A recent study combined subliminal and supraliminal priming (Damen et al., 2014). Result showed that congruency effects were similar to those discussed earlier for subliminal priming (see section 1.2.2.3), but were reversed when primes were supraliminal. This would be consistent with other suggestions that effort could enhance SoA (Demanet, Muhle-Karbe, Lynn, Blotenberg, & Brass, 2013; Lafargue & Franck, 2009). However, this study only used free choice trials, and the effects of priming on action were minimal. The authors (Damen et al., 2014) instead suggested that actions that were consciously biased were associated with a reduction in perceived freedom, and thus a reduction in SoA. In line with an influence of choice on SoA, increasing the response space available to choose from, i.e. from 1 to 7 buttons, has been associated with an increase in SoA (Barlas & Obhi, 2013). Nonetheless, “choice overload” effects have also been shown, in which having too many response options (e.g. 24) has a negative effect on motivation, behaviour, and preference judgements (Iyengar & Lepper, 2000). These findings highlight that

choice, as well as fluency vs. effort can have different effects on SoA depending on context, but these contextual effects remain poorly understood.

## **1.5. Aims of the Present Thesis**

The present thesis aimed to clarify the cognitive and neural processes underlying prospective contributions to SoA, related to the metacognitive monitoring of action selection processes. Thus, we aimed to go beyond the current centrality of explicit judgements of agency attribution in situations of social ambiguity. We did this by focussing on the cognitive processes that precede action, rather than by focussing on the consequences of action, as in previous studies. In addition, we used intuitively valid, but relatively novel measures of agency, namely explicit ratings of the feeling of control over an outcome. With these two key developments, we present a systematic family of several behavioural and electrophysiological experiments in healthy volunteers, designed to explore whether and how prospective cognitive processes contribute to the sense of agency.

In Chapter 2, we manipulated action selection with supraliminal stimuli across 3 experiments, to investigate the impact of awareness of conflict stimuli on the of selection fluency on SoA. We additionally investigated the influence of having a choice in what to do, or following instructions; and the effects of the timing of conflict. These studies also allowed us to test the generalisability of selection fluency effects on SoA by using a different manipulation of action selection.

Chapter 3 developed new manipulations of action selection in a more ecological setting, a videogame-like task. Across 3 experiments, different stages of action selection were manipulated: processing fluency, ambiguity in stimulus categorisation, response conflict. This dynamic setting also allowed us to probe the interaction between prospective and retrospective cues to SoA. Action selection manipulations were factorially combined with introducing a discrepancy between movements and visual feedback.

Electrophysiological correlates of prospective cues to SoA were investigated in Chapter 4. The subliminal priming paradigm was combined with a manipulation of whether participants could choose which action to make, or had to follow an instruction. We investigated a possible relation between SoA and pre- and/or post-decisional conflict monitoring ERPs, i.e. target-locked N2, and action-locked CRN. Additionally, we considered an ERP index of outcome monitoring, the feedback-related negativity. This allowed an exploration of whether selection fluency effects on SoA were mediated by changes in outcome processing.

Chapter 5 reports a multi-study analysis of the role of action selection to SoA, combining the studies described in Chapters 2 & 4 with other available studies (for a total of 7). This investigated the interaction between prospective, action selection related, vs. retrospective, outcome-based, cues to SoA. In particular, we assessed whether selection fluency effects might change over time, during action-outcome learning.

Chapter 6 investigated the influence of social context, more specifically diffusion of responsibility, on action selection and SoA. For this, we developed a task based on a helping scenario, where there was no ambiguity about outcome attribution. In addition to explicit agency ratings, we measured the FRN in outcome-locked ERPs. This allowed us to test whether diffusion of responsibility might be associated with online changes in SoA and outcome monitoring.

Finally, there will be a general discussion aiming to integrate the findings of the studies presented here and previous literature. In particular, we will discuss the relation between SoA and metacognition, and provide a tentative model of the contribution of metacognitive signals to different varieties of agency experience.







## **Chapter 2. Difficult Action Decisions Reduce the Sense of Agency: A Study Using the Eriksen Flanker Task**

*Previous research on prospective contributions to the sense of agency has mostly used subliminal priming to manipulate action selection. It remained unclear whether affecting action selection with supraliminal stimuli would have similar effects. Here, we used supraliminal flankers to manipulate action selection in response to a central target. Experiment 1 revealed that conflict in action selection, induced by incongruent flankers and targets, led to lower agency ratings over action outcomes, relative to neutral and congruent flanker conditions. Experiment 2 replicated this result, and extended it to free choice between alternative actions. Finally, Experiment 3 varied the stimulus onset asynchrony (SOA) between flankers and target. Action selection performance varied with SOA. Agency ratings were always lower in incongruent than congruent trials, and this effect did not vary across SOAs. Sense of agency is influenced by a signal that tracks conflict in action selection, regardless of the visibility of stimuli inducing conflict, whether choosing freely or following instructions, and even when the timing of the stimuli means that the conflict may not affect performance.*

A version of this Chapter can be found in following publication: Sidarus, N., & Haggard, P. (2016). Difficult action decisions reduce the sense of agency: A study using the Eriksen flanker task. *Acta Psychologica*, 166, 1–11. <http://doi.org/10.1016/j.actpsy.2016.03.003>

## 2.1. Introduction

Much research on the sense of agency has focused on the process of retrospectively comparing expected and actual action outcomes (Blakemore et al., 2002; Wegner & Wheatley, 1999). Recent studies have revealed that a metacognitive signal about the fluency of action selection can also contribute to the sense of agency prospectively (for a review, see (Chambon, Sidarus, et al., 2014). These studies used subliminal priming to manipulate action selection in an agency task (Chambon & Haggard, 2012; Chambon, Moore, et al., 2014; Chambon et al., 2013; Sidarus et al., 2013; Wenke et al., 2010). Participants make left or right actions according to a target arrow, which are followed by coloured circles – the action outcomes. Participants are then asked to judge how much control they felt over these circles. Unbeknownst to the subject, a small arrow – a prime – is briefly flashed before the target. When the prime is congruent with the target, and points in the same direction, action selection is easy; but when the prime is incongruent with the target, and points in the opposite direction, action selection is impaired, leading to slower reaction times (RTs) and more errors (e.g. Wenke et al., 2010). Results showed that the sense of agency over action outcomes was higher following congruently primed actions, compared to incongruently primed actions.

Importantly, outcomes could not be predicted by the action or the prime alone, but depended on the *congruency* between prime and target. Further, the effects of action selection on sense of agency could not be explained by participants relying on a retrospective monitoring of RTs. Tellingly, a further experiment manipulated the timing of stimuli to induce either a normal priming effect or a “negative compatibility effect” (NCE; Eimer & Schlaghecken, 1998). In the NCE, congruent primes impair rather than facilitate motor performance. This manipulation reversed the effects of primes on RTs, as expected, but judgements of agency were always higher for congruent priming, in both normal and NCE priming (Chambon & Haggard, 2012). The authors proposed a model in which the very initial action intention, triggered by the prime, could be compared with the executed action. Congruency between the initial intention and action would facilitate a metacognitive signal about action selection, and thus lead to a higher sense of agency. The later motor inhibitory

processes that caused NCE would occur downstream of this metacognitive readout of initial intention.

Since these primes were subliminal, participants were not aware that selection fluency was manipulated, and could not strategically decide to use fluency as a cue to agency (Chambon & Haggard, 2012; Wenke et al., 2010). Fluency can be thought of as a continuum between easy, or fluent, perceptual or cognitive processing, to effortful, or dysfluent, processing (Alter & Oppenheimer, 2009). Response conflict is an instance of highly effortful processing (Botvinick & Braver, 2015). Although the experience of selection fluency/dysfluency may be relatively weak, people may have a sense of “something going right/wrong” in congruent or incongruent trials respectively, without being able to identify why they have this feeling (Chambon, Sidarus, et al., 2014; Pacherie, 2008). It has been shown that people can reliably introspect on their experience of ease/difficulty in action selection, using a similar subliminal priming task (Desender et al., 2014), as well as with conflicting supraliminal stimuli (Morsella et al., 2009). This feeling could then become associated with subsequent events, such as action outcomes (Fritz & Dreisbach, 2013; Winkielman et al., 2015). Interestingly, similar effects are found when measuring agency at the end of a trial (e.g. Chambon & Haggard, 2012) and at the end of a block (Wenke et al., 2010). This suggests that the association between fluency experiences and outcomes could build up over time. Alternatively, the learning of action-outcome relations may be disrupted by dysfluent action selection.

In fact, the studies that used subliminal priming to manipulate selection fluency (e.g. Chambon et al., 2013; Wenke et al., 2010) differ considerably from previous research on the sense of agency, as they are focused on the *instrumental* learning of the relation between specific actions and a number of possible outcomes (Chambon, Filevich, et al., 2014). From this perspective, expertise with a given environment leads to a growing sense of ease, or flow, in selecting an action, which becomes associated with more predictable outcomes. On the other hand, research on the sense of agency has often focused on the *attribution* of agency. In such studies, action-outcome associations are often well known (Elsner et al., 2002), and may be

violated (Kühn et al., 2011), and/or there may be ambiguity about “who” caused a specific outcome, i.e. me vs. another agent (e.g. Wegner & Wheatley, 1999).

Response conflict induced by conscious stimuli has been shown to lead to a reduced sense of agency over one’s *actions* (Morsella et al., 2009). However, it remains unclear whether conscious stimuli that influence action selection might also alter the sense of agency over *action outcomes*. One suggestive study set out to manipulate the visibility of primes, while measuring judgements of agency over outcomes (Damen et al., 2014). Participants were aware of some primes, but not others. Prime words (“left” vs. “right”) were presented for a short or long duration, producing subliminal or supraliminal priming, respectively. Participants freely chose whether to press a left or right key once the following mask disappeared. Their action triggered a high or low tone after a variable delay, and participants judged their agency over the tone. For the subliminal priming condition, judgements of agency followed the pattern previously reported, i.e. higher ratings for trials in which the action was congruent with the prime, relative to prime-incongruent actions. However, for supraliminal primes, the effects were reversed, and higher ratings were found for prime-incongruent actions. The authors argued that awareness that one’s choice might have been biased by external input would reduce one’s sense of freedom and, in turn, one’s sense of agency.

Importantly, Damen et al.’s (2014) study showed effects of priming on the sense of agency, despite showing little or no effect of either subliminal or supraliminal primes on reaction times. Priming of choices was only found for supraliminal primes, in one of two experiments. Thus, there is little evidence that primes influenced action selection processes in their study. This contrasts with previous reports in which even subliminal primes reliably biased free choices (Kiesel et al., 2006; Klapp & Haas, 2005; Klapp & Hinkley, 2002; O’Connor & Neill, 2011; Schlaghecken & Eimer, 2004; Wenke et al., 2010). Instead, Damen et al. (2014) argued that action primes might influence agency judgements independently of influencing action selection, by affecting higher-order, conceptual representations of action and agency.

The present study aimed to clarify the contribution of action selection processes to sense of agency, using supraliminal stimuli to manipulate action selection across 3 experiments. To additionally test the generalizability of these effects, a novel task was used – the Eriksen flanker task (Eriksen & Eriksen, 1974). This is widely used to induce response conflict, and assess cognitive control dynamics. The flanker task was adapted and combined with the design from the aforementioned subliminal priming studies (Chambon & Haggard, 2012; Wenke et al., 2010). Participants responded according to a target letter (e.g. left for S, right for H), which could appear flanked by congruent (e.g. HHHHH) or incongruent flankers (e.g. SSHSS). A coloured circle appeared after a variable delay, and participants judged their control over that colour. In the incongruent flanker condition, the presence of flankers associated with the alternative action should lead to response conflict, and thus an increase in RTs and errors.

Experiment 1 aimed primarily to test how supraliminal stimuli relevant to action selection would affect the sense of agency in a situation where each action could produce one of a number of outcomes. Damen et al.'s results might suggest that the highest sense of agency would be found in the incongruent condition, when participants had to overcome conscious response conflict. However, if selection fluency has a general effect on the sense of agency then the highest sense of agency should be found in the congruent flanker condition. Additionally, we included a neutral condition, with task-irrelevant flankers (i.e. OOHOO) to try to distinguish facilitation and conflict effects on action (Kopp et al., 1996; Mansfield, van der Molen, Falkenstein, & van Boxtel, 2013; Taylor, 1977), and on the sense of agency. Finally, some previous studies measured agency ratings at the end of each trial, while others measured agency ratings at the end of a block. In this study, we exploratorily tested half of the participants with each method, though we did not have any strong prediction about interactions involving rating method.

Importantly, free vs. instructed choice could modulate how awareness of priming stimuli would influence the sense of agency. For subliminal priming, having a higher or lower proportion of free choice trials, relative to forced choice, did not interact with

the effects of action selection on agency (Wenke et al., 2010). However, this may be different for conscious priming. A participant who consciously perceives a prime might recruit cognitive control resources to resist its influence, potentially increasing their sense of agency. This possibility was assessed in Experiment 2. Forced choice (i.e. instructed) trials were randomly intermixed with free choice trials. A task-irrelevant target letter indicated a free choice trial, and appeared surrounded by task-relevant flankers (e.g. HHOHH). Hence, actions could be congruent or incongruent with the flankers, whether the action was instructed by the central, attended stimulus, or was endogenously chosen.

Additionally, the timing of stimuli affecting action selection, and thus response conflict, could be important. A sufficient amount of time may be needed between the appearance of biasing information and an instruction/go-signal to develop a clear awareness that one is either following or going against that information. One might then come to have a stronger sense of agency for overcoming external biases. Similarly, if there is enough time, cognitive control processes can inhibit the automatic motor activation induced by primes or flankers, thus abolishing their effects on motor performance (Flowers, 1990; Wascher, Reinhard, Wauschkuhn, & Verleger, 1999). In this case, choosing to go against the prime does not require any additional effort over choosing to go with the prime. Nonetheless, awareness of an external suggestion could still influence one's sense of agency.

To test the impact of the timing of conflicting stimuli, Experiment 3 parametrically varied the stimulus onset asynchrony (SOA) between flankers and target. Flankers could precede the target by 500 ms (-500 SOA) or 100 ms (-100 SOA), be simultaneous with the target (0 SOA), or follow the target after 100 ms (+100 SOA). Maximal congruency effects on performance are found for -100 and 0 SOA conditions, but only small or no effects are found for the -500 and +100 SOA conditions (Eriksen & Schultz, 1979; Flowers, 1990; Taylor, 1977; Wascher, Reinhard, Wauschkuhn, & Verleger, 1999; Willemsen, Hoormann, Hohnsbein, & Falkenstein, 2004). We hypothesized that the -500 SOA condition would allow sufficient time for suppression of the flankers, and potentially alter effects of conflict



on sense of agency. The -100 SOA condition was expected to still show important effects on action selection, but the clear precedence of the flankers to the target might alter the subjective experience of conflict and agency. The 0 SOA condition should replicate our previous effects. In addition, the +100 SOA condition would serve to assess whether the temporal precedence of flankers or target might influence agency processing. If congruency between a *first* intention and the action performed is the important comparison for agency, as suggested by Chambon & Haggard (2012), then this condition should not affect agency even if it showed minor effects on performance. Since choice did not interact with fluency effects on agency in Experiment 2, only forced choice trials were used.

## **2.2. Experiment 1**

### **2.2.1. Materials and Methods**

#### **2.2.1.1. Participants**

The study was approved by the UCL Research Ethics Committee. Twenty-five participants (13 female, *mean* = 23.62, *SD* = 3.98) were recruited, based on an *a priori* power calculation. For this, we used previous reports of prime compatibility on agency in ratings in operant reaction-time tasks (Chambon et al., 2013), since no previous study to our knowledge had investigated flanker congruency effects on sense of agency over action outcomes. With a Cohen's  $d_z$  of 0.66 (Chambon et al., 2013), power = 0.8, and alpha = 0.05, a minimum sample size of 21 was indicated, but a slightly larger number were recruited, in anticipation of possible attrition. Participants gave written informed consent to participate in the study and received payment of £7.5/h. All were right-handed, with normal or corrected-to-normal vision, did not suffer from colour blindness, and had no history of psychiatric or neurological disorders. There were two groups of participants: odd-numbered participants rated agency on every trial, while even-numbered participants rated agency at the end of each block. One participant in the block-wise rating group was excluded due to difficulties in distinguishing outcome colours.

#### **2.2.1.2. Apparatus and Materials**

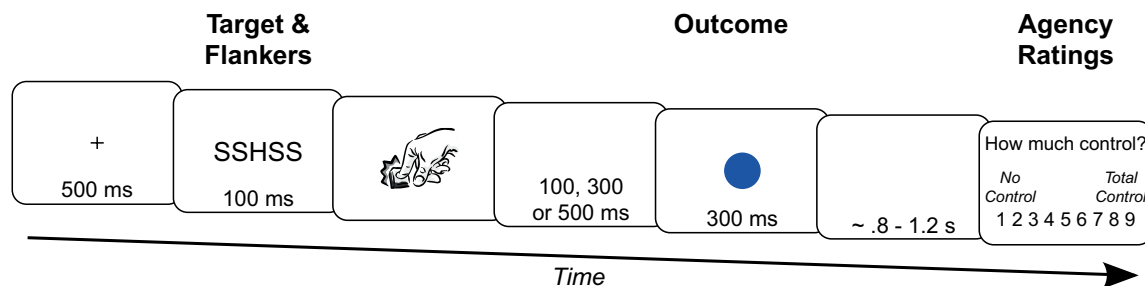
Participants were seated approximately 50 cm from a computer screen. The experiment was programmed and stimuli delivered with Psychophysics Toolbox v3 (Brainard, 1997; Kleiner, Brainard, & Pelli, 2007; Pelli, 1997), running on Matlab (MATLAB 8.1, The MathWorks Inc., Natick, MA, 2013). During a trial, stimuli were presented in a mono-spaced font, Lucida Console. A fixation cross was presented in 18 point font size. Target letters consisted of S's or H's, while flankers consisted of S's, H's or O's. These were presented in 30 point font size, with the 5 letter array subtending 3.2° visual angle. Participants responded by pressing one of two keys on a keyboard. Outcome stimuli consisted of a circle of 2.8° presented in one of 6 colours (red, blue, green, yellow, orange and pink). Different colours were used in the training phase.

All participants gave agency ratings on a 9-point Likert scale. The trial-wise ratings group completed the rating procedure on the computer. For the block-wise ratings, participants were first asked to rank order the coloured circles (cut-outs) on a sheet of paper, and then gave a Likert rating for each colour.

#### **2.2.1.3. Design and Procedure**

The task involved making actions in response to targets, which were surrounded by distracting flankers. The action triggered the appearance of a coloured circle – the action outcome. Participants were instructed to pay attention to the relation between their actions and the outcomes that followed, as they were required to judge these relations at the end of each trial or each block, for the respective group. Participants had to respond with a left or right key press according to a central target letter (S or H, respectively). The assignment of target letters to a left or right action was counterbalanced across participants. Participants were instructed to ignore the flankers and focus on the central letter. Flankers could be congruent with the central target – e.g. HHHHH, and thus with the required action; incongruent – e.g. SSHSS; or neutral – e.g. OOHOO (Matchock & Mordkoff, 2007; Taylor, 1977). Flanker-target congruency was randomly varied across trials.

Outcome colours were dependent on both the congruency condition and the action performed. Thus, each action (left vs. right) was associated with three outcomes, one for each congruency condition (cf. Wenke et al., 2010). The condition-to-colour mapping varied across the blocks, so participants had to learn the action-outcome relations anew in each block, and were informed of this. The six outcome colours were rotated in a Latin square across the 6 blocks, and the block mapping was randomised. Each colour appeared once in each experimental condition, thus cancelling out any idiosyncratic colour preferences. To ensure that the frequency of each coloured outcome was equal despite differences in error rates across flanker-action congruency conditions, error trials were replaced at the end of a block. Additionally, the action-outcome interval was varied orthogonally to the congruency factor. This was not a variable of interest, but served as a dummy variable, ensuring that participants were exposed to a range of experiences, varying from low sense of agency (for delayed outcomes) to high sense of agency (for less delayed outcomes; Haggard, Clark, & Kalogeras, 2002; Wenke et al., 2010).



**Figure 2.1.** Timeline of an example incongruent trial, with trial-wise ratings. Participants responded according to a central target letter, surrounded by distractors. This triggered the appearance of a coloured circle, after a variable delay. Participants gave agency ratings at the end of each trial, for the trial-wise rating group; or completed a ranking/rating procedure at the end of each block, for the block-wise rating group.

Participants were asked to judge how much control they felt over the coloured circles that were triggered by their actions (Chambon & Haggard, 2012; Wenke et al., 2010). For the trial-wise rating group, a 9-point Likert scale was presented at the end of each trial, where 1 was labelled “No Control” and 9 was labelled “Total Control”. The block-wise ratings group completed a ranking and rating procedure on a paper sheet

at the end of each block. Participants were instructed to rank order coloured circles on the sheet across 6 rankings, from “Most Control” to “Least Control”. After ranking, participants gave a rating of their sense of control on the Likert scale described above.

The study started with a training block of 24 trials, to allow participants to get acquainted with the experiment and the agency ratings procedure. Participants were given a chance to ask questions and repeat the training if desired. To avoid colour mapping repetitions, different colours were used during the training and experimental phases. At the end of the study, participants completed a short debriefing questionnaire.

### **2.2.1.4. Timeline**

Each trial started with a fixation cross presented for 500 ms. The flankers and target array appeared for 100 ms (Gratton et al., 1992; Rodríguez-Fornells, Kurzbuch, & Münte, 2002). Participants responded to the target within a 1.2 second window. If the response was correct, an outcome colour followed the response after a variable delay of 100, 300 or 500 ms. Outcome duration was 300 ms. If an incorrect response or no response was given, a black cross was presented for 300 ms. For the trial-wise rating group, the agency rating scale appeared after 800 to 1200 ms, and remained on the screen until a response was given. For both groups, the inter-trial interval varied randomly between 1 and 1.5 seconds. Each block consisted of 72 trials, and there were 6 blocks overall. At the end of each block, the block-wise rating group completed the ranking/rating procedure. All participants were allowed to take short breaks between blocks.

### **2.2.1.5. Data Analysis**

For the block-wise ratings group, rating sheets were coded and the data computerised. Any blocks where mistakes were made in the ranking/rating procedure were excluded from analysis. Mistakes could involve mismatches between the ranking and rating, or the repetition of a colour name. This resulted in the exclusion of 1 block in 2 participants, and 2 blocks in another participant.

Reaction times (RTs), error rates and agency ratings were submitted to a 2 x 3 mixed-design analyses of variance (ANOVA). The between-subjects factor was group: trial- or block- wise ratings group; and the within-subjects factor was flanker-action congruency: congruent, neutral or incongruent. Planned comparisons were used to test differences between congruency levels. For the block-wise ratings group, agency ranks were submitted to a Friedman's non-parametric test to assess the main effect of flanker-action congruency. Wilcoxon pairwise tests were used for planned comparisons. Within subjects 95% confidence intervals were obtained for the main effect of congruency (Loftus & Masson, 1994).

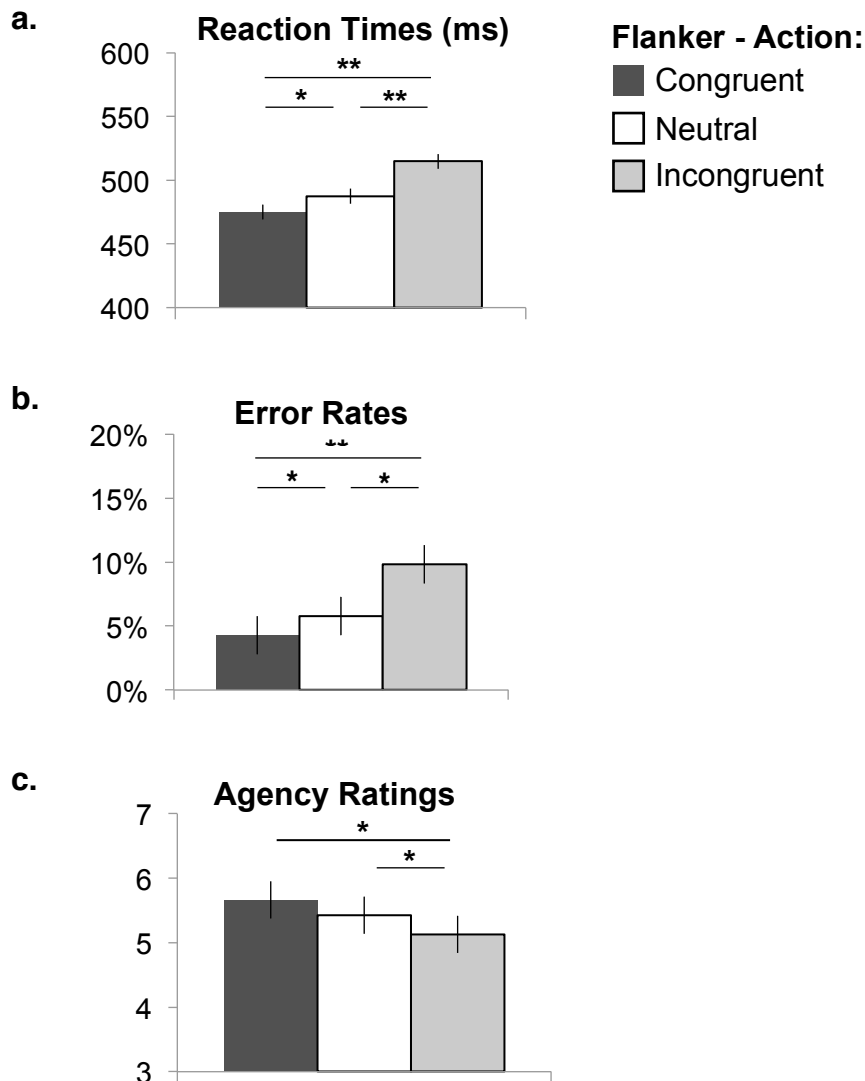
## 2.2.2. Results

### 2.2.2.1. Action Selection

Analyses of RTs showed a significant effect of flanker-action congruency ( $F_{(2, 44)} = 64.46$ ,  $p < .001$ ,  $\eta_p^2 = .75$ ; see **Figure 2.2.a**), but no effect of group nor interaction ( $F_s < 1$ ). Planned comparisons revealed that RTs were significantly slower ( $ps < .001$ ) in the incongruent condition ( $mean = 514.78$ ,  $SD = 67.84$ ) compared to the neutral ( $mean = 487.42$ ,  $SD = 70.99$ ) and congruent conditions ( $mean = 475.02$ ,  $SD = 65.25$ ). RTs were also significantly slower in the neutral compared to the congruent condition ( $p = .004$ ).

Analyses of error rates revealed a significant main effect of congruency ( $F_{(2, 44)} = 18.55$ ,  $p < .001$ ,  $\eta_p^2 = .46$ , Greenhouse-Geiser correction; see **Figure 2.2.b**). Planned comparisons showed that participants made significantly more errors in the incongruent ( $mean = 9.82\%$ ,  $SD = 8.39\%$ ) compared to neutral ( $mean = 5.79\%$ ,  $SD = 5.65\%$ ; incongruent vs. neutral:  $p = .001$ ), and congruent conditions ( $mean = 4.29\%$ ,  $SD = 4.79\%$ ; incongruent vs. congruent:  $p < .001$ ). The neutral condition also led to significantly more errors than the congruent condition ( $p = .017$ ). Additionally, there was a significant main effect of group ( $F_{(1, 22)} = 5.73$ ,  $p = .026$ ,  $\eta_p^2 = .21$ ), as the trial-wise ratings group made significantly more errors than the block-wise ratings group. This presumably reflects higher task difficulty for the trial-wise rating group, as they had to give agency ratings in each trial, which meant they

had to press different keys. In contrast, the block-wise rating group could focus exclusively on responding to the target, and could keep their fingers on the response keys throughout a block. Finally, there was no significant interaction between group and congruency ( $F_{(1, 22)} = 2.65$ ,  $p = .10$ ,  $\eta_p^2 = .11$ , Greenhouse-Geiser correction). However, our study may not have had strong enough power to investigate interactions involving the effect of group.



**Figure 2.2.** Results for Experiment 1. Panel **a.** shows mean reaction times across flanker-action congruency conditions (collapsed across groups), and **b.** shows mean error rates. Both facilitation and conflict effects can be seen in RTs and error rates. **c.** Mean agency ratings show only an effect of conflict, such that agency ratings were significantly reduced following incongruent relative to neutral and congruent trials. Error bars show the within subjects 95% confidence intervals for the main effect of congruency.

\*  $p < 0.05$ , \*\*  $p < 0.001$

### 2.2.2.2. Agency Ratings

The ANOVA on agency ratings revealed a significant main effect of congruency ( $F_{(2, 44)} = 4.70$ ,  $p = .014$ ,  $\eta_p^2 = .18$ ; see **Figure 2.2.c**). Planned comparisons confirmed that the incongruent condition ( $mean = 5.13$ ,  $SD = 1.57$ ) led to significantly lower ratings compared to the congruent ( $mean = 5.66$ ,  $SD = 1.74$ ; incongruent vs. congruent:  $p = .013$ ), and the neutral condition ( $mean = 5.42$ ,  $SD = 1.63$ ; incongruent vs. neutral:  $p = .039$ ), whereas the congruent and neutral conditions were not significantly different ( $p = .21$ ). There was no significant effect of group ( $F_{(2, 44)} = 1.29$ ,  $p = .29$ ,  $\eta_p^2 = .013$ ), nor a significant group x congruency interaction ( $F_{(1, 22)} = 0.30$ ,  $p = .59$ ,  $\eta_p^2 = .055$ ).

For the block-wise group, agency ranks were also analysed, and results showed a significant main effect of congruency ( $\chi^2(2) = 8.73$ ,  $p = .013$ ). Planned comparisons replicated the pattern of results seen for the agency ratings: the incongruent condition ( $median = 3.25$ ,  $SD = 0.99$ ) led to significantly lower agency ranks than the congruent condition ( $median = 4.00$ ,  $SD = 0.66$ ; incongruent vs. congruent:  $Z = -2.57$ ,  $p = .010$ ,  $r = -0.37$ ), and the neutral condition ( $median = 3.50$ ,  $SD = 0.50$ ; incongruent vs. neutral:  $Z = -2.27$ ,  $p = .024$ ,  $r = -0.33$ ); whereas there was no significant difference between congruent and neutral conditions ( $Z = -0.99$ ,  $p = .32$ ,  $r = -0.14$ ).

### 2.2.3. Discussion

Experiment 1 showed that flanker-action congruency influenced action selection as predicted. The sense of agency over action outcomes was significantly reduced following dysfluent action selection, compared to fluent selection. This replicates recent work demonstrating a prospective contribution of action selection processes to the sense of agency (Chambon & Haggard, 2012; Chambon et al., 2013; Sidarus et al., 2013; Wenke et al., 2010), and generalises the finding across different behavioural tasks. So far, most studies used subliminal priming to manipulate action selection (Chambon, Sidarus, et al., 2014), or assessed agency over the *action* (Morsella et al., 2009). To the best of our knowledge, the present study is the first to

show a reduction in the sense of agency over *action outcomes* following dysfluent action selection, even though participants could consciously perceive the stimuli that influenced action selection.

Previous studies (Chambon & Haggard, 2012; Chambon et al., 2013; Sidarus et al., 2013; Wenke et al., 2010) used subliminal priming to manipulate action selection in order to preclude the explicit awareness that one's action was manipulated. Additionally, this increased uncertainty about the outcomes, since they were contingent on both the action and the congruency between the (invisible) prime and the action. That is, as the primes were not consciously perceived, the relation between prime-action congruency and specific outcomes could not be represented, hence outcomes were never fully predictable. In contrast, as participants were aware of the flankers in the present study, they could learn the full contingency schedule between the letter strings and outcome colours. For example, in a given block, participants could learn that the letter array "SSSSS" was followed by a green circle, whereas "HSHHH" was followed by a red circle. Debriefing confirmed that most participants were aware of this relation. Moreover, the causes of difficulties in action selection, i.e. incongruent flankers, were now clearly available to participants. Nevertheless, the same effects of action selection fluency on agency ratings were found, irrespective of perceptual awareness of the stimulus trigger.

Moreover, there was no significant difference in the fluency effects on agency across the two rating procedures, i.e. trial- vs. block- wise ratings. While the same effects had been shown using both procedures, this was the first study to combine them. Previous studies suggest that action selection fluency affects agency online (Chambon, Moore, et al., 2014; Chambon et al., 2013). Additionally, the association between different fluency experiences and ensuing outcomes can be retained in memory, at least for long enough to accumulate over the course of a block of trials, as seen here and in Wenke et al. (2010).

The inclusion of a neutral condition allowed us to distinguish an enhanced sense of agency due to facilitation of action selection, from a reduction of agency due to response conflict. Only the effect of conflict in action selection yielded a significant



modulation of agency ratings (see **Figure 2.2**). When flankers were congruent with the central target, participants were faster and made fewer errors, than when the flankers were neutral. Additionally, incongruent flankers led to significantly slower RTs and more errors, compared to neutral flankers. However, while agency ratings were significantly lower following incongruent flankers, compared to neutral and congruent flankers, the trend for higher ratings following congruent compared to neutral flankers was not statistically significant.

It should be noted that other baseline conditions, and different tasks, could yield a different pattern of facilitation/conflict (Jonides & Mack, 1984). The present study used task-irrelevant stimuli as neutral flankers, which yielded both facilitation and conflict effects on performance. As congruency effects on agency ratings are smaller than for RTs, the absence of a facilitation effect could result from a lack of statistical power within-subjects. Additionally, the between-subjects design resulted in a smaller sample in each group. Thus, we had low statistical power for between-subjects effects and interactions. However, these between-subjects effects did not form the focus of our predictions. These considerations mean that null between-subjects effects should be interpreted with particular care. Nevertheless, our key results, of congruency effects on agency ratings, are based on within-subjects comparisons. Further, they are consistent with those obtained with the subliminal priming paradigm (Chambon & Haggard, 2012). There, the reduction in agency ratings following incongruent, compared to neutral primes, was larger than the increase in ratings following congruent primes, though neither was statistically significant. A positive sense of agency may be a “default state” (Blakemore et al., 2002; Sidarus et al., 2013). Reduced agency may be triggered by disruptions in the intention-action-outcome chain, which may produce a salient experience relevant to agency judgement (Chambon, Sidarus, et al., 2014).

Our results contrast sharply with those of Damen et al. (2014). That study reported higher agency ratings when participants chose an action incongruent with a supraliminal prime, compared to when they chose a prime-congruent action. Importantly, free choice trials were used in their study, whereas here participants had

to follow the instruction of a central flanker. Experiment 2, therefore, investigated whether choice may interact with the effects of flanker congruency on sense of agency, when biasing stimuli are consciously perceived. Free and forced choice targets were randomly intermixed, such that actions could be congruent or incongruent with the flankers, whether the action was instructed by the central, attended stimulus, or was endogenously chosen.

### **2.3. Experiment 2**

#### **2.3.1. Materials and Methods**

##### **2.3.1.1. Participants**

Participant recruitment and study approval was as in Experiment 1. Twenty-four participants were tested (13 female, *mean* = 21.50, *SD* = 3.02).

##### **2.3.1.2. Design and Procedure**

Testing conditions and stimuli were the same as in Experiment 1, except that instead of a neutral flanker condition, the letter O now served as a neutral target in free choice trials. In free choice trials, the neutral target was surrounded by flankers associated with a left or right action. For example, if the array “SSOSS” was presented, participants could choose whether to act congruently with the flankers and make a left action, or act incongruently with the flankers and choose a right action. Thus, flanker-action congruency was not related to the stimuli, but rather reflected the participants’ action choice. In forced choice trials, the congruent or incongruent conditions were as described in Experiment 1. The new 2 (choice: free vs. forced) x 2 (congruency: congruent vs. incongruent) design meant that 8 outcome colours were used, 4 associated with each hand, 1 per choice x congruency condition. The colours were Latin square rotated across 8 blocks of 64 trials, and the condition-colour block mappings were randomised.

All participants gave agency ratings at the end of each trial, thus the trial timeline was the same as the trial-wise group in Experiment 1. Only 2 action-outcome

intervals were used (200 and 400 ms), to reduce the overall number of conditions. As in Experiment 1, the study began with a training block of 32 trials, and ended with a debriefing questionnaire.

### 2.3.1.3. Data Analysis

Reaction times were submitted to a 2 x 2 ANOVA, with choice (free vs. forced) and flanker-action congruency (congruent vs. incongruent) as within-subjects factors. Agency ratings were submitted to a similar ANOVA, with action-outcome interval (200 vs. 400 ms) as an additional within-subjects factor. For free choice trials, the proportion of flanker congruent choices was analysed with a one-sample t-test against a .05 chance level. For forced choice trials, error rates were analysed with a paired-samples t-test comparing congruent and incongruent conditions. Within subjects 95% confidence intervals for pairwise comparisons were calculated separately for free and forced choice trials (Pfister & Janczyk, 2013).

## 2.3.2. Results

### 2.3.2.1. Action Selection

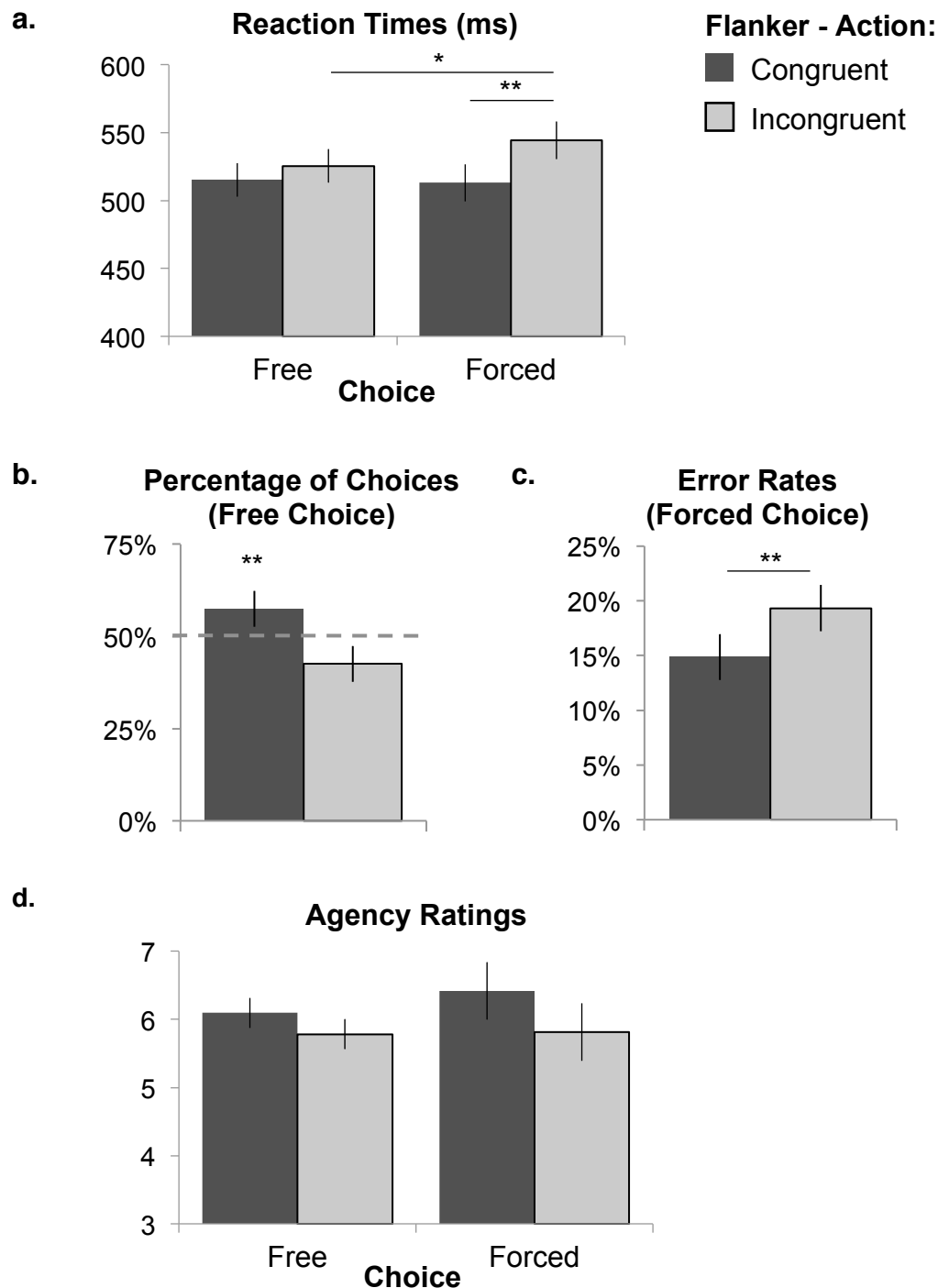
Analyses of RTs revealed no significant main effect of choice ( $F_{(1, 23)} = 1.65$ ,  $p = .21$ ,  $\eta_p^2 = .067$ ), a significant main effect of congruency ( $F_{(1, 23)} = 20.76$ ,  $p < .001$ ,  $\eta_p^2 = .47$ ; see **Figure 2.3.a**), and a significant choice x congruency interaction ( $F_{(1, 23)} = 5.67$ ,  $p = .026$ ,  $\eta_p^2 = .20$ ). Simple effects t-tests showed a significant congruency effect for forced choice trials, i.e. slower RTs for the incongruent ( $mean = 544.28$ ,  $SD = 88.73$ ) than the congruent condition ( $mean = 513.09$ ,  $SD = 83.49$ ), and a similar modest trend for free choice trials (free congruent:  $mean = 515.27$ ,  $SD = 88.85$ ; free incongruent:  $mean = 525.50$ ,  $SD = 96.85$ ; one-tailed, free:  $t_{(23)} = -1.72$ ,  $p = .050$ , Cohen's  $d_z = -.35$ ; forced:  $t_{(23)} = -4.68$ ,  $p < .001$ , Cohen's  $d_z = .96$ ). Additionally, incongruent trials led to significantly slower RTs in forced compared to free choice ( $t_{(23)} = -2.18$ ,  $p = .040$ , Cohen's  $d_z = .44$ ). Choice did not affect RTs in congruent trials ( $t < 1$ ).

In free choice trials, flanker congruent choices were made in 57.47% ( $SD = 5.72$ ) of trials (see **Figure 2.3.b**). A one sample t-test showed that the proportion of flanker-congruent choices was significantly different from chance ( $t_{(23)} = 6.40$ ,  $p < .001$ , Cohen's  $d_z = 1.31$ ). For forced choice trials, a paired samples t-test on error rates showed that the incongruent condition ( $mean = 19.33\%$ ,  $SD = 11.73$ ) led to significantly more errors than the congruent condition ( $mean = 14.85\%$ ,  $SD = 10.15$ ;  $t_{(23)} = -4.39$ ,  $p < .001$ , Cohen's  $d_z = -.90$ ; see **Figure 2.3.c**).

### 2.3.2.2. Agency Ratings

An ANOVA on agency ratings revealed a significant main effect of congruency ( $F_{(1, 23)} = 12.70$ ,  $p = .002$ ,  $\eta_p^2 = .36$ ). Flanker-incongruent actions ( $mean = 6.25$ ,  $SD = 1.07$ ) led to lower agency ratings than flanker-congruent actions ( $mean = 5.80$ ,  $SD = 1.23$ ; see **Figure 2.3.d**). Critically, there was no significant main effect of choice ( $F_{(1, 23)} = 1.48$ ,  $p = .24$ ,  $\eta_p^2 = .061$ ), nor a significant choice by congruency interaction ( $F_{(1, 23)} = 2.32$ ,  $p = .14$ ,  $\eta_p^2 = .092$ ).

There was a marginal effect of action-outcome interval ( $F_{(1, 23)} = 3.65$ ,  $p = .069$ ,  $\eta_p^2 = .14$ ), such that ratings for the long interval (400ms;  $mean = 6.08$ ,  $SD = 1.10$ ) were higher than for the short interval (200ms;  $mean = 6.01$ ,  $SD = 1.12$ ). These results are inconsistent with previous findings using other tasks (Chambon & Haggard, 2012; Chambon, Moore, et al., 2014; Haggard et al., 2002; Sidarus et al., 2013). In previous studies, using a wider range of intervals, higher ratings were found for shorter intervals, recalling Hume's concept of temporal contiguity as a cue for causation (Hume, 1740). Importantly, action-outcome interval did not interact with the factors of interest – choice and congruency ( $F_s < 1$ ). Since action-outcome interval was not a factor of interest, this factor will not be discussed further.



**Figure 2.3.** Results for Experiment 2. **a.** Mean reaction times for free and forced choice trials, and flanker-action congruency conditions. Congruency effects were larger for forced choice trials, and RTs in incongruent trials were slower in forced choice conditions. **b.** Percentage of trials in which participants chose the action that was congruent or incongruent with the flankers in free choice trials, revealing a bias towards flanker-congruent choices. **c.** Mean error rates, in flanker-congruent or incongruent actions, for forced choice trials, with more errors being made in incongruent trials. **d.** Mean agency ratings across conditions, showing only a main effect of flanker-action congruency, with lower ratings following flanker-incongruent actions, for both free and forced choice trials. Error bars display the pairwise within subjects 95% confidence intervals, calculated separately for free and forced choice trials. \*  $p < 0.05$ , \*\*  $p < 0.001$

### 2.3.3. Discussion

Experiment 2 showed that action selection was influenced by flanker – action congruency in both free and forced choice trials. Flankers biased choice, such that participants were ~ 7 % more likely to ‘freely’ select actions corresponding to the flanker suggestion, compared to against it. Similar biases have been found using subliminal priming (Mattler & Palmer, 2012; Schlaghecken & Eimer, 2004; Wenke et al., 2010). Flanker-incongruent actions led to significantly slower RTs in forced choice trials, with a similar trend in free choice trials. Additionally, incongruent forced choice trials led to significantly slower RTs than incongruent free choice trials. Hence, the cost on performance of freely choosing an action incongruent with the flankers was smaller than the cost of following an instruction with incongruent flankers. Consistently, a greater flexibility for changes of mind has been shown for free, compared to forced, choices (Fleming, Mars, Gladwin, & Haggard, 2009). Crucially, response conflict, induced by supraliminal flankers, significantly reduced the sense of agency over action outcomes for both instructed and freely chosen actions.

Our results additionally show that the discrepancy between our findings and those of Damen et al. (2014) cannot be explained by whether participants could freely choose which action to perform, or had to follow an instruction. Although null effects should be interpreted with care, the absence of an interaction between choice and congruency seen here is consistent with a previous subliminal priming study (Wenke et al., 2010). In Wenke et al.’s (2010) study, free and forced choice trials were intermixed, and free choices were effectively biased by subliminal primes, similarly to our results. On the other hand, Damen et al. (2014) found little effect of sub- or supraliminal primes on choice, possibly due to the exclusive use of free choice trials. This could have allowed participants to decide which action to make before the beginning of a trial, and thus before the prime was presented. In fact, it has been shown that priming effects seen in blocks of intermixed free and forced choice trials are abolished in blocks with only free choice trials (Klapp & Haas, 2005; Schlaghecken & Eimer, 2004).

Nonetheless, Damen et al. (2014) did find priming effects on agency. The authors argued that the observed reduction in the sense of agency when following a conscious prime could have been due to a reduced sense of freedom. Using only free choice trials could have potentially increased the overall sense of freedom experienced in the task, relative to mixed conditions, rendering a reduction in that perceived freedom, due to conscious biases, more salient. This sense of freedom may affect agency at a higher, conceptual level, and independently of action selection.

Another relevant difference between the two studies, which is related to action selection, lies in stimulus timing. In Damen et al., the prime preceded the go signal by 250 ms in the supraliminal priming condition, and there was no time limit for response. In contrast, in our study, flankers and targets were presented simultaneously, speed was emphasised, and a tight response window was imposed. Hence, a 'sufficient' amount of time may be necessary for a realisation that one's actions are being biased, and thus override the normal relation between selection fluency and sense of agency. To assess whether the timing of conflict stimuli may influence the sense of agency, the interval between flankers and target onset was parametrically varied in Experiment 3.

## **2.4. Experiment 3**

### **2.4.1. Materials and Methods**

#### **2.4.1.1. Participants**

Participant recruitment and study approval was as in Experiments 1 & 2. Twenty-six participants were tested (13 female, mean age = 23.08,  $SD = 3.63$ ). One participant was excluded as she did not follow instructions, and sometimes used only one hand to press the left and right key.

### 2.4.1.2. Design and Procedure

Testing conditions were the same as in Experiment 2, but with only forced choice trials. Additionally, the flanker-target stimulus onset asynchrony (SOA) was randomly varied across the trials. Flankers could appear: 500 ms before target onset (-500 SOA); 100 ms before target onset (-100 SOA), simultaneously with the target (0 SOA); or 100 ms after the target (+100 SOA). To accommodate the varying SOA conditions, target duration was now set to 150 ms (Wascher et al., 1999). Flankers were displayed until the target duration elapsed. Action-outcome intervals were also changed to 100 and 500 ms to enhance the discriminability of the 2 intervals, while keeping the experimental session short.

Each block included 4 outcome colours, one per action x congruency condition, orthogonal to the flanker-target SOA conditions. To obtain a similar number of trials per SOA x congruency condition to the previous experiments, 12 blocks of 64 trials were used. To ensure that each outcome colour appeared only once for each action x congruency condition, 12 colours were used overall in the experiment. These were rotated with a Latin square across the 12 blocks, in groups of 4, and the block mappings were randomised. The 12 colours were shown to participants at the beginning of the study to confirm that they could reliably distinguish them. Participants were also instructed that the colours or the relation between action and colours could change across blocks, so they needed to learn them anew in each block. As in the previous experiments, the study began with a training block of 32 trials, and ended with a debriefing questionnaire.

### 2.4.1.3. Data Analysis

RTs and error rates were submitted to a 4 x 2 repeated measures ANOVA with the factors flanker-target SOA (-500, -100, 0, +100) and flanker-action congruency (congruent vs. incongruent). Agency ratings were submitted to a similar ANOVA that additionally included the factor action-outcome interval (100 vs. 500 ms). Greenhouse-Geisser corrections were used whenever the sphericity assumption was violated. Bonferroni adjusted post-hoc tests were used to probe the main effect of SOA. The SOA x congruency interactions were investigated with paired samples



t-tests, with a Bonferroni adjustment, to test congruency effects across SOAs. Within subjects 95% confidence intervals for the pairwise differences between congruency conditions were calculated separately for each SOA (Pfister & Janczyk, 2013).

## 2.4.2. Results

### 2.4.2.1. Action Selection

Analyses of RTs revealed significant main effects of SOA ( $F_{(3, 72)} = 240.77, p < .001, \eta_p^2 = .91$ ), and congruency ( $F_{(1, 24)} = 60.40, p < .001, \eta_p^2 = .72$ ), and a significant SOA x congruency interaction ( $F_{(1, 24)} = 9.28, p < .001, \eta_p^2 = .28$ ). Post-hoc tests to explore the main effect of SOA showed that all pairwise comparisons between SOAs were significant ( $ps < .001$ ). As **Figure 2.4.a** shows, RTs were faster with earlier presentation of the flankers. Probing the SOA x congruency interaction revealed significant congruency effects at each SOA ( $ps < .001$ ), except at -500 SOA ( $t_{(1, 24)} = -1.04, p = .31$ ).

Analyses of error rates showed no significant effect of SOA ( $F_{(1, 24)} = 1.08, p = .36, \eta_p^2 = .04$ ), a significant main effect of congruency ( $F_{(1, 24)} = 31.61, p < .001, \eta_p^2 = .57$ ), and a significant SOA x congruency interaction ( $F_{(1, 24)} = 5.01, p = .003, \eta_p^2 = .17$ ). Post hoc tests revealed significant congruency effects for -100 and 0 SOA (-100 SOA:  $t_{(1, 24)} = -5.08, p < .001$ , Cohen's  $d_z = -1.02$ ; 0 SOA:  $t_{(1, 24)} = -3.54, p = .002$ , Cohen's  $d_z = .71$ ), but not for -500 or +100 SOA (-500:  $t_{(1, 24)} = -.39, p = .70$ , Cohen's  $d_z = -0.078$ ; +100:  $t_{(1, 24)} = -1.58, p = .13$ , Cohen's  $d_z = -0.32$ ; see **Figure 2.4.b**).



**Figure 2.4.** Results for Experiment 3. **a.** Mean reaction times across flanker-target stimulus onset asynchronies (SOAs), and flanker-action congruency conditions. There are congruency effects at all SOAs except -500 ms. **b.** Mean error rates across condition, showing a larger congruency effects at -100 and 0 SOAs. **c.** Mean agency ratings across conditions. There was only a main effect of congruency, with incongruent trials leading to lower ratings than congruent trials, regardless of flanker-target SOA. Error bars show the within subjects 95% confidence intervals for the congruency pairwise differences, for each SOA. \* -  $p < .05$ , \*\* -  $p < .001$

### 2.4.2.2. Agency Ratings

Analyses of agency ratings revealed a marginal main effect of congruency ( $F_{(1, 24)} = 3.99$ ,  $p = .057$ ,  $\eta_p^2 = .14$ ), in the predicted direction: incongruent flankers ( $mean = 6.42$ ,  $SD = 1.57$ ) led to lower ratings compared to congruent flankers ( $mean = 6.67$ ,  $SD = 1.47$ ; see **Figure 2.4.c**). Notably, there was no main effect of SOA ( $F_{(3, 72)} = .87$ ,  $p = .46$ ,  $\eta_p^2 = .035$ ), and no interaction between SOA and congruency ( $F_{(3, 72)} = .40$ ,  $p = .75$ ,  $\eta_p^2 = .017$ ). The absence of SOA effects on agency ratings can be clearly observed in **Figure 2.4.c**.

Finally, there was a trend towards a main effect of action-outcome interval ( $F_{(1, 24)} = 3.27$ ,  $p = .083$ ,  $\eta_p^2 = .12$ ), with long intervals (500 ms;  $mean = 6.57$ ,  $SD = 1.47$ ) leading to higher agency ratings than short intervals (100 ms;  $mean = 6.53$ ,  $SD = 1.50$ ). There was also a marginal interaction between congruency and action-outcome interval ( $F_{(1, 24)} = 3.48$ ,  $p = .074$ ,  $\eta_p^2 = .13$ ), which was not a focus of prediction, and so was not explored further. The remaining interactions were not significant ( $ps > .18$ ). Both action-outcome interval results are inconsistent with previous priming studies (Chambon & Haggard, 2012; Chambon, Moore, et al., 2014; Sidarus et al., 2013). Even though the difference between the two intervals was increased, relative to Exp. 2, varying the flanker-target SOA may have changed the perception of the subsequent action-outcome interval, and disrupted its normal effects on agency. Since action-outcome interval was not a manipulation of interest, this will not be discussed further.

### 2.4.3. Discussion

Results showed that flanker effects on action selection were modulated by the flanker-target SOA. As predicted, flankers had no effect on action selection at -500 SOA, but incongruent flankers did lead to performance costs with the other SOAs (see **Figure 2.4.a** and **b**). Additionally, there was a gradual increase in RTs with increasing SOA, possibly due to an alerting effect of early flankers, also found in previous studies (Flowers, 1990; Taylor, 1977; Wascher et al., 1999; Willemsen et al., 2004). Critically, there was no significant interaction between flanker-target SOA

and congruency on agency ratings. That is, incongruent conditions led to (marginally) lower agency ratings than congruent conditions, but did so similarly across flanker-target SOAs (see **Figure 2.4.c**), including SOAs where flankers had no performance effects.

These results are inconsistent with the hypothesis outlined above of an interaction between the timing of conflict during action selection and the direction of fluency effects on agency. That hypothesis suggested that SOAs favouring successful inhibitory cognitive control might lead to higher agency ratings for incongruent, rather than congruent flankers. At -500 SOA, we found efficient inhibitory cognitive control, resulting in no congruency effect on RTs or error rates, yet sense of agency was still higher for congruent than incongruent trials. Therefore, the results of Damen et al. (2014) cannot be explained by a longer time delay between a biasing influence and action allowing the recruitment of cognitive control to efficiently overcome those biases.

The dissociation seen here between congruency effects on motor performance and on agency ratings is, however, consistent with Damen et al. (2014), where priming influenced agency but not action selection. The authors argued that the effects were independent of selection fluency, but rather due to priming of conceptual representations of action, or to influencing the experience of freedom. A dissociation between motor effects and agency was also found in a subliminal priming study, using NCE priming (Chambon & Haggard, 2012). It was proposed that congruency between an initial prime's suggestion and the executed action could serve as a fluency signal that would increase the sense of agency.

However, neither of these proposals can fully account for our results, since they would predict that only congruency between the *first* intention and the action should matter. Our results show that the appearance of incongruent flankers 100 ms *after* the target still affected the sense of agency, even though the action performed remained congruent with the first intention, which was presumably triggered by the target. Therefore, it seems that holding conflicting intentions is key for the observed reduction in the sense of agency, rather than the precise dynamics of the selection

process. Importantly, this condition still led to congruency effects on motor performance, consistent with earlier reports (Eriksen & Schultz, 1979; Taylor, 1977). Action selection processes take time, and will be susceptible to disruptions occurring within a given time window. When using arrow stimuli in the flanker task, no performance effects were found with a +100 SOA (Wascher et al., 1999). Thus, the window in which action selection can be disrupted may vary depending on whether the stimulus is imperative in nature.

Our results are compatible with a view of the sense of agency as resulting from an integration of information about conflict over a wider time-window than the time-window of action selection. It has been argued that fluency/conflict signals are relatively non-specific with respect to their sources, and have only a general influence (Winkielman et al., 2015). The temporal sensitivity of such signals, and of their integration in the sense of agency, may be low relative to the precise temporal dynamics of action selection and execution. To better characterise this window of temporal integration, future studies could include more flanker-target asynchrony values. In particular, one might ask whether flankers continue to influence the sense of agency even when presented so late that they no longer influence reaction times.

### **2.5. General Discussion**

Overall, our results suggest that the sense of agency over an action outcome is informed by cognitive processes occurring prior to action execution, particularly those processes involved in initiating a correct rather than an inappropriate action. In many situations, action control requires identifying an appropriate target, and then selecting and initiating the corresponding action, while avoiding the influence of distractors. The feeling of control over the consequences of action is influenced by these processes. Part of the content of agency judgements appears to derive from monitoring processes that detect response conflict during action selection.

Interestingly, we found that sense of agency was insensitive to the specific dynamics of conflict at the level of motor performance. Thus, the prospective, premotor signals that influence sense of agency appear to signal a disruption in action selection

whenever conflict emerges, regardless of whether the conflict is successfully resolved, and of how performance is affected (cf. Chambon & Haggard, 2012). Additionally, this putative monitoring system can integrate information about action selection in a time window that is broader than that which affects selection at a motor level. Moreover, the effects of action selection on the sense of agency can be independent of the effects of choice, and of the effects of being aware of influences on one's action or choice. That is, regardless of whether we have a choice in what to do, and whether we are aware of stimuli that could bias our decisions, dysfluent or difficult action selection can lead to a reduction in our sense of agency over action outcomes. Finally, we have shown that these effects generalise across tasks.

Our results imply that the sense of agency depends on some internal signal related to selecting between alternative actions. In that regard, our results are compatible with 'metacognitive' theories of agency (Metcalfe & Greene, 2007). Where might these internal signals be found within the motor system? The supplementary motor area (SMA) is necessary for triggering the automatic inhibition processes thought to underlie NCE priming, whereas upstream regions such as the pre-SMA are not (Sumner et al., 2007). Such automatic inhibition processes were not found to disrupt the sense of agency (Chambon & Haggard, 2012). The pre-SMA has in turn been implicated in monitoring response conflict, elicited both by conscious and unconscious stimuli (van Gaal, Scholte, Lamme, Fahrenfort, & Ridderinkhof, 2010). Relatedly, the premotor cortex, but not the primary motor area, has been shown to contribute to metacognitive judgements of perceptual confidence (Fleming et al., 2014). More specific to the present findings, an fMRI study (Chambon et al., 2013) used the subliminal priming paradigm to study congruency effects on the sense of agency. This study showed that the dorsolateral pre-frontal cortex was sensitive to response conflict, and was associated with the angular gyrus, wherein higher activity was linked to a greater reduction in agency ratings. Together, these studies suggest that the metacognitive monitoring of action selection that informs the sense of agency, may rely on higher-order action representations in premotor and prefrontal areas, rather than low-level motor signals in the primary motor cortex.

Importantly, the congruency effects on agency seen here are not due to a retrospective inferential process, but rely on prospective signals from action monitoring processes. As the flankers were clearly visible, one might be tempted to think that the observed effects could result from a retrospective comparison between the flankers and the target, or action, namely at a conceptual level. However, this would imply that neutral flankers would lead to a loss of agency, as they were visibly different from the target. Instead, the effects seen here appear specifically related to *conflict* in action selection. Experiment 1 showed no significant difference between congruent and neutral flankers, but only a significant reduction in agency following incongruent flankers. Although such null effects should be interpreted with care, especially due to potentially low statistical power, they suggest that a perceptual or conceptual mismatch may not be sufficient to explain our results. Rather, an incongruent action plan should be triggered at some stage, for a reduced sense of agency. In fact, subliminal priming was used in previous studies to manipulate action selection but preclude such post-hoc, conceptual inferences. This method showed a consistent trend for a larger cost of conflict on agency ratings than a facilitation effect (Chambon & Haggard, 2012). Our Experiment 3 is also consistent with a prospective account: the presence of conflicting motor plans during the trial led to a loss of agency, even when the interval between flankers and target was sufficient to resolve the conflict. Consistently, supplementary analyses showed that congruency effects on the sense of agency could not be fully explained by RT monitoring (see **Appendix A**, also cf. Chambon & Haggard, 2012). The subjective experience of conflict may linger, even after the motor conflict has been resolved. Conflict signals are especially motivationally significant since they can indicate a need to adjust subsequent behaviour (Botvinick & Braver, 2015; Holroyd & Yeung, 2012). As such, they may have a greater impact on the sense of agency than fluency experiences. Additionally, a positive sense of agency may be a ‘default’, and thus we are especially sensitive to disruptions to the normal flow of voluntary action (Chambon, Sidarus, et al., 2014).

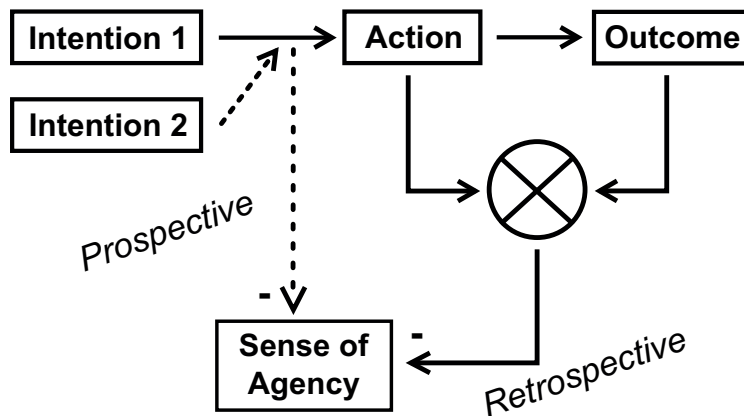
Our results clearly contrast with some reports that effort or difficulty can *enhance* sense of agency (Damen et al., 2014; Demanet, De Baene, Arrington, & Brass,

2013). Why, then, do effort and conflict sometimes increase sense of agency, and sometimes reduce it? The relation between fluency or effort and the sense of agency is complex and remains poorly understood (Nahmias, 2005; Pacherie, 2008). Often when intentional actions unfold without any obstacles, the sense of fluency can result in a strong sense of agency, as “*everything went according to plan*”. Yet, effort can also enhance the sense of agency. When a need for cognitive control can be anticipated, some proactive conflict processing (Braver, 2012) may become part of the action plan. This may highlight the sense of self, and of being engaged with task at hand. In contrast, when disruptions are unexpected, executive control will be triggered reactively by conflict signals. We speculate that these two sources of cognitive control may have different effects on sense of agency. In particular, proactively embedding effort into the action plan may be associated with an increase in the sense of agency (*I knew it would be tricky, but I managed it*), however, the unexpected or unwanted need for added effort could instead lead to a reduction in our sense of agency (*suddenly I had to deal with all these things*).

In addition, the context or the framing of a task could modulate how conflict influences agency. In Damen et al.’s study, each action triggered a specific outcome (a beep with a given pitch) after a variable delay (0-600 ms). Participants were instructed that sometimes they would cause the beep to occur (the outcome), but other times it would be caused by the computer. Thus, the task and the agency question were framed in terms of *attributing* the cause of the outcome to the self, or to another. Also, subliminal and supraliminal priming were randomised, so participants presumably experienced wide variations in degree of influence from the primes. In contrast, our studies focused on the *instrumental* aspect of agency, as participants were asked to judge the *strength* of the relation between various actions and outcomes, rather than invoking alternative agents. That is, our study focussed on ‘concomitant variation’ between a single agent’s different instrumental actions and their outcomes, rather than on attribution of outcomes to agents. Both processes are relevant to agency, but conflict between alternative actions might have different effects on each of them. Further research is needed to clarify the conditions under which conflict can enhance, rather than reduce, the sense of agency.



Our results are consistent with previous proposals that the sense of agency integrates information from multiple sources (Synofzik et al., 2008), and over time (Chambon, Sidarus, et al., 2014; Farrer et al., 2013). In addition to retrospective processes related to outcome monitoring, there is also a prospective component related to action selection (see **Figure 2.5**). Action selection monitoring can detect conflicting intentions and prospectively signal a loss of agency. After this, outcome monitoring can assess action outcome intervals and outcome identity for a mismatch with predictions or expectations, and retrospectively signal a loss of agency. If the smooth flow between intention – action – outcome remains unperturbed, the sense of agency can remain at a default level. Additionally, higher-order beliefs and contextual information can also influence the sense of agency (Moore & Fletcher, 2012; Synofzik et al., 2008). We found that choice, awareness of biases and timing of conflict did not interact with the effects of selection fluency. However, they may make independent contributions to the sense of agency, depending on context, or other cues.



**Figure 2.5.** Prospective and retrospective contributions to the sense of agency. The sense of agency is prospectively informed by monitoring action selection. When this action monitoring system detects an intention that conflicts with the to-be-executed intention, it sends a signal indicating a loss of agency. Once the action outcome is known, this can be compared with a prediction of the outcome, based on the executed action. When there is a mismatch between the predicted and actual outcomes, an outcome monitoring system can retrospectively signal a loss of agency. If the normal flow from intention, to action, to outcome is disrupted, the sense of agency is reduced.

### **2.5.1. Conclusions**

Across the experiments reported here, the sense of agency was prospectively informed by monitoring the processes of action selection. When conflicting intentions were present, the sense of agency over action outcomes was reduced. The effect of conflict on the sense of agency was independent of awareness of the causes of conflict, of free vs. instructed action selection, and of the timing of conflicting information during action selection. Finally, these effects generalised across tasks, from subliminal priming of actions, to the Eriksen flanker task, thus revealing a new approach for further investigating prospective contributions to the sense of agency.

These findings support the view that the sense of agency is especially sensitive to a disruption in the normal flow of intentional action, from an intention or goal to its corresponding action, to the desired/expected consequences (Chambon, Sidarus, et al., 2014; Haggard & Chambon, 2012). Importantly, fluency of action selection was independent of the actual statistical contingency between actions and outcomes in these experiments. Selection fluency does not guarantee successful agency: one can know exactly what to do, and still fail to produce an intended outcome. However, selection fluency may serve as a useful heuristic to guide our sense of agency, as it often predicts successful outcomes (Haggard & Chambon, 2012). Prospective agency processes based on action selection may thus help to bridge the time gap between action and outcome.





## Chapter 3. New Avenues for Investigating the Prospective Sense of Agency

*The present study focused on the prospective influence of action selection processes to the sense of agency in an ecologically valid, dynamic setting. Across three experiments, different manipulations were used to target different stages of action selection: visual processing fluency, categorisation ambiguity and response conflict. Additionally, we probed the relative contributions of prospective, action selection-based cues, and retrospective, outcome-based cues to the sense of agency. Manipulations of action selection were combined with discrepant visual feedback of action on some trials, thus manipulating proximal outcome monitoring.*

*Results show that fluency in action selection influenced the sense of agency across tasks. Moreover, discrepant visual feedback led to a large reduction in the sense of agency, which tended to predominate over manipulations of action selection. The sense of agency appears to be highly sensitive to disruptions of motor control but, under successful motor control, disruptions at varying stages of action selection have a robust effect on the sense of agency.*

### 3.1. Introduction

As we interact with the world around us, our experience is typically coloured by a sense of agency, a feeling that we are in control of our actions and, through them, can control events in the outside world (Haggard & Tsakiris, 2009). Much research has focused on how actions are linked to their outcomes. This has shown that the sense of agency depends on a retrospective comparison between expected or desired action outcomes and actual outcomes (e.g. Blakemore, Wolpert, & Frith, 2002; Wegner & Wheatley, 1999). Moreover, the importance of linking intentions and actions has been recently highlighted (for a review, see Chambon, Sidarus, et al., 2014). In fact, modern theoretical frameworks emphasise that the sense of agency results from the integration of multiple cues (Moore & Fletcher, 2012; Synofzik et al., 2008), which may become available at different times (Farrer et al., 2013; Haggard & Chambon, 2012). Moreover, metacognitive processes are involved in evaluating the output of action and outcome monitoring systems (Haggard & Chambon, 2012; Metcalfe & Greene, 2007; but see Chambon, Filevich, & Haggard, 2014).

Recent studies have used subliminal priming of actions to manipulate the fluency of action selection, in a simple paradigm in which participants respond according to directional arrows and trigger the appearance of coloured circles (Chambon & Haggard, 2012; Chambon, Moore, et al., 2014; Chambon et al., 2013; Sidarus et al., 2013; Wenke et al., 2010). Participants report a stronger sense of agency over action outcomes when primes induce fluent, compared to dysfluent, action selection. Therefore, a metacognitive signal about the fluency of action selection processes contributes to the sense of agency prospectively, and long before the outcome is known.

Interestingly, it has been shown that judgements of agency (JoAs) are influenced by the metacognitive monitoring of performance in a game, but are still highly sensitive to actual disruptions of control (Metcalfe & Greene, 2007). In these studies, a computer game was used in which participants move a mouse cursor (a box) along a horizontal bar to catch falling Xs, while avoiding Os. Introducing a discrepancy between one's mouse movements and the visual feedback of the cursor movements

(i.e. proximal outcomes) leads to a reduction in both performance and judgements of performance (JoPs), but an even greater reduction in JoAs. Additionally, when Xs or Os disappear without being touched (i.e. distal outcomes), JoAs are reduced, independently of how performance was affected. Nonetheless, the reduction in JoAs associated with disruptions of proximal outcomes is larger than the disruption of distal outcomes (cf. Metcalfe, Eich, & Miele, 2013). Together, these results suggest that the sense of agency is preferentially tuned to monitor proximal cues, tied to action and motor control, and indeed reflects the statistical contingency between actions and outcomes.

The present study aimed to further investigate the role of prospective, action selection-based, cues in the sense of agency. To this end, different manipulations of action selection were employed, and adapted to the computer game described above (Metcalfe & Greene, 2007). This allowed us to test the generalisability of the effects of action selection across manipulations, but also in a dynamic environment, with greater ecological validity than the experimental paradigms typically used to study the sense of agency. Additionally, we aimed to compare the relative contribution of prospective cues to agency with that of retrospective, outcome-based cues by manipulating proximal outcomes. Manipulations of action selection were thus combined with manipulating the discrepancy between action and visual feedback.

In this game, an action consists of moving the cursor to the horizontal location of a target (i.e. an X). The expected proximal outcome is that the cursor will move to the location matching one's mouse movements. The expected distal outcome is that the target will disappear once caught. This means action selection is determined by: first, detecting a stimulus; second, categorising it as a target or distractor; and, third, deciding whether to move the box towards it (for targets), while also avoiding distractor stimuli (i.e. O's). Therefore, with some changes to the stimuli, we could manipulate these three different stages of action selection.

Many studies have shown that fluency in visual processing or decision making can affect a variety of judgements, such as confidence, liking, or familiarity (for a review,

see Alter & Oppenheimer, 2009). Previous studies, which used action priming, argued that action selection fluency influences the sense of agency (e.g. Wenke et al., 2010). Priming can influence processing fluency, in addition to inducing response conflict, but, it remains unclear whether fluency in perceptual processing alone could have a similar effect on the sense of agency. To test this, in Experiment 1, we manipulated stimulus processing fluency through visual masking. This essentially disrupted the very first stage of the stimulus-response-outcome chain, namely, identifying and locating target stimuli. If fluency in action selection has a general effect on the sense of agency, we would predict that this manipulation would lead to a reduced sense of agency.

In Experiment 2, we manipulated the uncertainty associated with categorising stimuli as targets or distractors by varying stimulus ambiguity. When categorising highly ambiguous stimuli, uncertainty about the accuracy of the categorisation would be higher. It has been suggested that the consequences of uncertain decisions may be seen as less blame-worthy than the consequences of more informed decisions (Heath & Tversky, 1991). Thus, greater uncertainty during action selection could lead to a reduction in the sense of agency. Additionally, a large number of highly ambiguous stimuli would render action selection processes more difficult, as many items would be hard to categorise. This increased difficulty could also lead to a lower sense of agency.

Finally, in Experiment 3, we interfered with the later, pre-motor aspects of action selection, namely, deciding whether to approach or avoid a given stimulus. Response conflict was induced by placing incongruent flankers around central targets or distractors (Eriksen & Eriksen, 1974). For example, M could indicate targets and C indicate distractors. These items would appear surrounded by congruent (e.g. MMM) or incongruent (e.g. CMC) flankers. The simultaneous detection of a target and a distractor would suggest two conflicting responses: approach and avoid that spatial location. Detecting a distractor flanker could elicit an avoidant response, which would need to be overcome if it was flanking a target; while an approach response could be erroneously elicited by a target flanker placed



around a distractor. A condition in which only congruently flanked items appeared was compared to a condition in which most items were congruently flanked, and to a condition in which few items were congruently flanked. The degree of response conflict was expected to lead to a corresponding reduction in the sense of agency.

### **3.2. Experiment 1**

This experiment investigated the effects of visual processing fluency on the sense of agency, by adding a visual noise mask on the screen in some trials.

#### **3.2.1. Materials and Methods**

##### **3.2.1.1. Participants**

Twenty-three Columbia University or Barnard College students volunteered to participate for course credit, and gave written informed consent (12 female, mean age = 20.05, SD = 2.52, age not recorded for 1 person due to technical error). All were right-handed, with normal or corrected-to-normal vision, and neurologically healthy. The procedures described here conform to the guidelines of the APA concerning the protection of human subjects, and were approved by the Columbia Internal Review Board.

##### **3.2.1.2. Apparatus**

The experiments were conducted on iMac computers, using a mouse on mouse pad. The programme was developed using custom-built scripts running on Python. White Xs and Os were presented on a grey background. Each trial started with 10 stimuli of each type (targets vs. distractors). In unmasked trials, the background was grey (115/255 RGB scale). In masked trials, the noise mask consisted of a Gaussian-filtered patch of randomly distributed greyscale intensities. This was applied on top of the main game screen, except for the half grey horizontal bar and white box (i.e. the mouse cursor), which were drawn on top of the mask. This enabled masking of the stimuli based on which participants decided *what* to do, while allowing participants to track their movements equally well in both masked and unmasked conditions. On

some trials, we also introduced a discrepancy between the participant's movements of the mouse, and the movements of the cursor on the screen, termed "turbulence". In the turbulence condition, the movement of the box depended on the following noise function:

$$\Delta x' = \Delta x + \sigma \sin(2 \pi t / 2.4)$$

where  $\Delta x'$  is the movement of the box on the screen,  $\Delta x$  is the distance the participant actually moved the mouse,  $t$  is time in seconds, and  $\sigma$  is the amplitude of the noise wave.

### 3.2.1.3. Design and Procedure

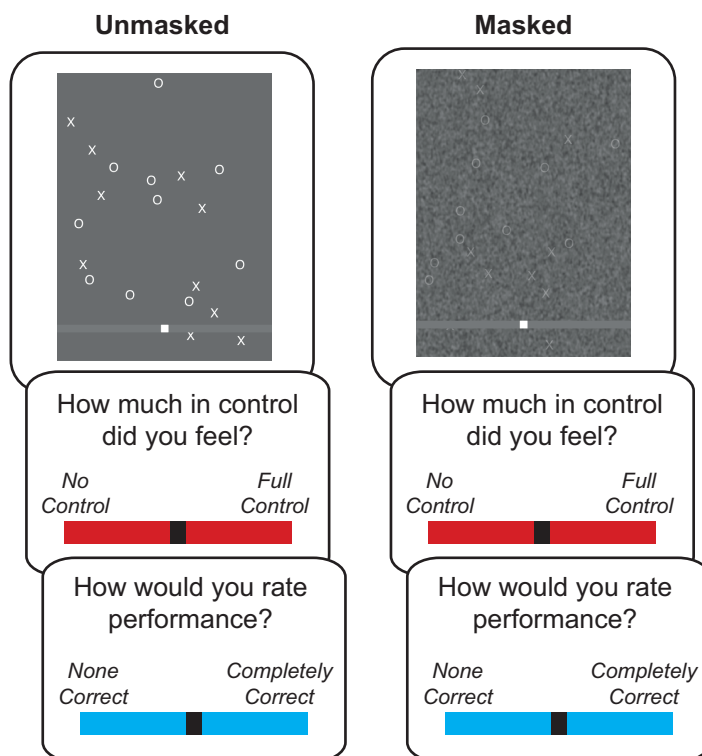
The basic procedure and instructions are described elsewhere (Metcalf & Greene, 2007). Briefly, participants played a game in which they observed Xs and Os scrolling down a screen, and moved a white box along a grey horizontal track with a mouse (see **Figure 3.1**). Participants were instructed to catch one letter with their white box (e.g. Xs), while avoiding the other letter (e.g. Os; counterbalanced between participants). Once caught, the items disappeared, and auditory feedback indicated whether a target or distractor was hit, with a ping or thud sound, respectively. If the items were not caught, they continued scrolling down to the bottom of the screen.

Participants played the game for 30 seconds, and then gave JoAs and JoPs about the game they had just played. For the JoAs, participants were asked to judge how much control they felt over the game, using a red visual analogue scale (VAS) ranging from "No Control" to "Full Control". For JoPs, participants were asked to rate their performance in the game using a blue VAS, ranging from "None Correct" to "Completely Correct". Participants moved a slider with the mouse, and pressed the space bar to select their rating.

To influence action selection, the fluency of visual processing was manipulated by either presenting a normal screen ("unmasked" trials), or adding a visual noise mask on top of the game ("masked" trials; see **Figure 3.1**). This masking made it harder to

detect the targets and distractors, thus rendering action selection more difficult. Additionally, to interfere with proximal outcome monitoring, the movements of the mouse cursor (the white box) were manipulated. In some trials, the cursor accurately followed the movement of the mouse (“no turbulence”). In other trials, a noise function was applied to the movements of the cursor (“turbulence”). These two manipulations, visual masking and cursor turbulence, were factorially combined, resulting in 4 trial (i.e. game) types. These were quasi-randomised across 6 blocks, such that each trial type was played once before the next block.

Before starting the experiment, participants played a training game, followed by JoAs and JoPs. They were given a chance to ask any questions, and either play another training game, or start the experiment. At the end of the experiment, participants answered a short questionnaire about the experiment and were debriefed.



**Figure 3.1.** Task outline for Experiment 1. Visual processing fluency was manipulated by adding a visual noise mask on some trials. After playing the game for 30 s, participants gave judgements of agency (on a red VAS), followed by judgements of performance (blue VAS).

### 3.2.1.4. Data Analysis

Performance in the game was assessed as a  $d'$  score from signal detection theory (Green & Swets, 1966), which measured discrimination between targets and distractors. The  $d'$  calculation was adjusted for instances of zero false alarms (Mill & O'Connor, 2014; Snodgrass & Corwin, 1988). JoAs and JoPs were quantified as a percentage of the VAS scale. Mean  $d'$ , JoAs and JoPs were submitted to repeated measures ANOVAs, with the factors masking (unmasked vs. masked) and turbulence (no turbulence vs. turbulence).

Additionally, we assessed whether any effects of visual masking on JoAs could be explained by a reduction in JoPs. For this, a hierarchical linear regression model (also known as linear mixed-effects models) was used to model single-trial level data. JoAs were modelled by the factors masking and turbulence, coded as unmasked = 0, masked = 1; no turbulence = 0, turbulence = 1. JoPs were added as a covariate, after standardising within participants. This analysis was conducted using the *lme4* package (Bates et al., 2014) in R (R Core Team, 2015). Parameter estimates ( $b$ ) and their associated t-tests ( $t$ ,  $p$ ), calculated using the Satterthwaite approximation for degrees of freedom (Kuznetsova, Brockhoff, & Christensen, 2015), are presented to show the magnitude of the effects, with bootstrapped 95% confidence intervals.

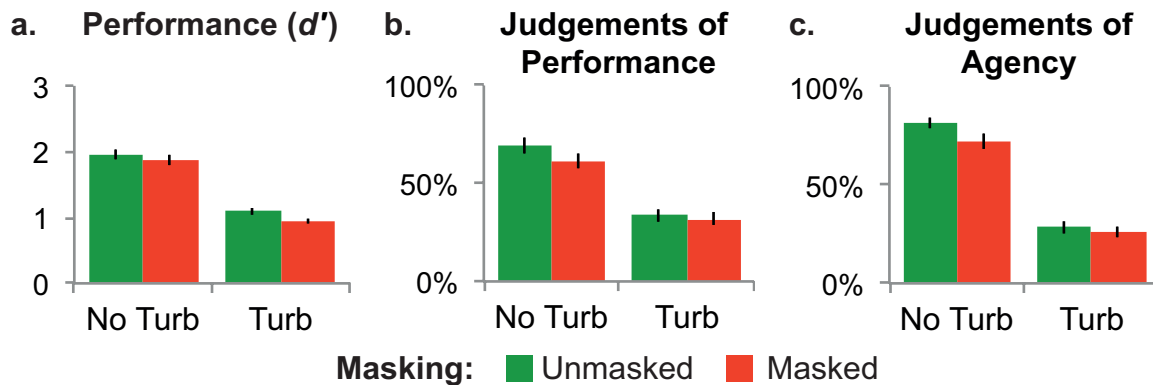
### 3.2.2. Results

Analysis of  $d'$  showed that discrimination between targets and distractors was significantly lower for masked, relative to unmasked, trials (mean difference = 0.12, SD = 0.15,  $F_{(1, 22)} = 13.67$ ,  $p = 0.001$ ,  $\eta_p^2 = 0.38$ ; see **Figure 3.2.a**). Turbulence also led to a significant reduction in  $d'$ , relative to no turbulence trials (mean diff. = 0.89, SD = 0.25,  $F_{(1, 22)} = 302.66$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.93$ ). There was no significant interaction between the factors ( $F_{(1, 22)} = 0.91$ ,  $p = 0.35$ ,  $\eta_p^2 = 0.040$ ).

Similarly, JoPs were significantly lower for masked, relative to unmasked, trials (mean diff. = 4.60%, SD = 4.73,  $F_{(1, 22)} = 20.66$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.48$ ; see **Figure**

**3.2.b).** Turbulence also led to a significant reduction in JoPs, relative to no turbulence trials (mean diff. = 32.67%, SD = 17.96,  $F_{(1, 22)} = 75.83$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.78$ ). There was a marginally significant interaction between the factors ( $F_{(1, 22)} = 3.53$ ,  $p = 0.074$ ,  $\eta_p^2 = 0.14$ ).

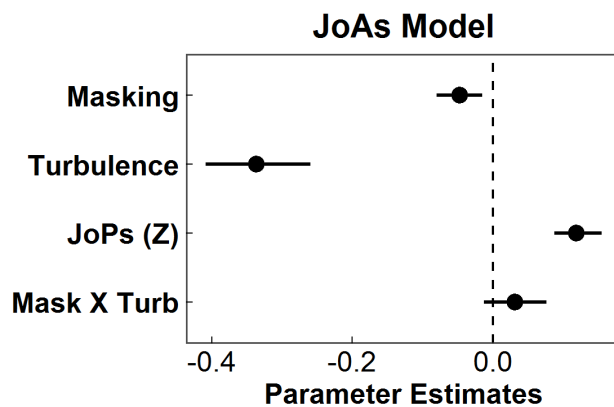
Results for JoAs (see **Figure 3.2.c**) showed a significant reduction for masked, relative to unmasked, trials (mean diff. = 6.13%, SD = 5.39,  $F_{(1, 22)} = 28.52$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.57$ ). Turbulence also led to a significant reduction in JoAs, relative to no turbulence trials (mean diff. = 49.58%, SD = 18.67,  $F_{(1, 22)} = 161.65$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.88$ ). Moreover, there was a significant interaction between the factors ( $F_{(1, 22)} = 5.81$ ,  $p = 0.025$ ,  $\eta_p^2 = 0.21$ ). Simple effects t-tests showed that, for no turbulence trials, masking led to a significant reduction in JoAs (mean diff. = 9.48%, SD = 10.65,  $t_{(22)} = 4.27$ ,  $p < 0.001$ ,  $d_z = 0.89$ ); whereas this reduction was only marginally significant for turbulence trials (mean diff. = 2.65%, SD = 6.19,  $t_{(22)} = 2.05$ ,  $p = 0.053$ ,  $d_z = 0.43$ ). Turbulence had a significant effect in both masking conditions (unmasked: mean diff. = 52.97%, SD = 19.91,  $t_{(22)} = 12.76$ ,  $p < 0.001$ ,  $d_z = 2.66$ ; masked: mean diff. = 46.14%, SD = 19.87,  $t_{(22)} = 11.14$ ,  $p < 0.001$ ,  $d_z = 2.32$ ).



**Figure 3.2.** Results of Experiment 1. Effects of masking and turbulence on mean  $d'$  (a.), JoPs (b.), and JoAs (c.). No Turb = no turbulence, Turb = turbulence. Error bars show the standard error of the mean.

Finally, we assessed whether the effect of masking on JoAs could be accounted for by changes in JoPs. Previous studies (e.g. Metcalfe, Eich, & Miele, 2013) showed that perceived performance is used as cue to agency, predicting a general, positive

relation between JoPs and JoAs. Therefore, JoAs were modelled by the experimental factors, and JoPs (standardised) were entered as a covariate (see **Figure 3.3**). The results revealed that JoPs were positively related to JoAs ( $b = 0.12$ ,  $t_{(22.56)} = 6.39$ ,  $p < 0.001$ , 95% CI = [0.087, 0.16]), as predicted. Importantly, masking remained a significant predictor of JoAs ( $b = -0.047$ ,  $t_{(21.53)} = -2.98$ ,  $p = 0.007$ , 95% CI = [-0.080, -0.015]), as did turbulence ( $b = -0.34$ ,  $t_{(20.27)} = -8.08$ ,  $p < 0.001$ , 95% CI = [-0.41, -0.26]). The interaction between masking and turbulence was no longer significant ( $b = 0.031$ ,  $t_{(51.91)} = 1.48$ ,  $p = 0.14$ , 95% CI = [-0.013, 0.076]).



**Figure 3.3.** JoAs model for Experiment 1. Parameter estimates, with bootstrapped 95% confidence intervals, for modelling JoAs by the factors masking and turbulence, and by JoPs (within-participants Z-score).

### 3.2.3. Discussion

In Experiment 1, visual masking was used to disrupt the processing fluency of the stimuli that drove participants' actions. This, in turn, disrupted the fluency of action selection, as it made the process of deciding which action to make, i.e. where to move the cursor, more difficult. The results showed that visual masking disrupted both objective and subjective measures of performance, as expected. Moreover, visual masking led to a reduction in JoAs, which could not be explained by perceived differences in performance, due to increased task difficulty. That is, the mere disruption of processing fluency led to a loss of agency.

These results are consistent with research showing that fluency can affect a number of metacognitive judgements (see Alter & Oppenheimer, 2009 for a review).

Moreover, previous studies have shown that dysfluent action selection is associated with a reduction in the sense of agency (Chambon, Sidarus, et al., 2014). These studies used response conflict to manipulate the fluency of action selection. Here, we show that disrupting action selection at the early stage of stimulus processing can lead to a reduction in the sense of agency.

In addition to visual processing, we manipulated whether the consequences of one's actions matched one's intentions, by introducing a discrepancy between the movement of the mouse and the movement of the cursor, termed turbulence. When the cursor on the screen did not accurately track the mouse's movements, there was a reduction in both objective and subjective measures of performance, as well as a large reduction in JoAs. This reduction in JoAs was independent of differences in JoPs, and was larger than the effect of turbulence on JoPs. These findings replicate previous studies that used a similar manipulation (e.g. Metcalfe, Eich, & Miele, 2013; Metcalfe & Greene, 2007).

Finally, while discrimination performance reflected additive effects of the visual masking and mouse turbulence manipulations, the effects of these factors on metacognitive judgements of agency were partially underadditive. This underadditivity can be seen in the significant interaction between visual masking and turbulence effects on average JoAs. Overall, visual masking had a minor effect on JoAs compared to the effect of turbulence. Additionally, we found that visual processing dysfluency especially disrupted the sense of agency when the cursor accurately followed the movements of the mouse, but less so when turbulence was introduced. On the other hand, turbulence had a large and robust effect across masking conditions. The larger effect of turbulence on performance suggests that it may be a more salient cue for agency than visual processing fluency. In fact, average JoPs also showed a somewhat underadditive effect of the two manipulations, possibly denoting that turbulence was a stronger cue for both metacognitive judgements.

Nonetheless, the interaction between masking and turbulence on JoAs was no longer significant after accounting for differences in JoPs, in a multi-level regression

model. We may speculate that the apparent underadditive effect is due to the difference in the size of the effects of masking and turbulence on JoAs. Future studies should attempt to balance the effects of masking and turbulence on performance, to clarify whether these underadditive effects on metacognitive judgements are linked to differences in the salience of the two manipulations. One may speculate that they are not, given that mouse turbulence is a reliable and direct indicator of a loss of agency. Participants were less able to implement their intended actions, and thus had objectively less control over the game. Visual masking only made action selection more difficult, but did not interfere with the ability to interact with the game. Alternatively, selection fluency may have a general effect on JoAs, consistent with other work on the effects fluency on metacognitive and affective judgements (Alter & Oppenheimer, 2009).

### **3.3. Experiment 2**

This experiment aimed to investigate whether the ease, or difficulty, of classifying stimuli as targets for action would influence sense of agency, by varying stimulus ambiguity.

#### **3.3.1. Materials and Methods**

##### **3.3.1.1. Participants**

Procedures were as described above. Twenty-four new Columbia University or Barnard College students (10 female, mean age = 20.79, SD = 2.93) volunteered to participate for course credit, and gave written informed consent. Five participants were left-handed, and the remaining were right-handed.

##### **3.3.1.2. Apparatus**

The basic setup was as in Experiment 1, except for the following. The target and distractor items now consisted of six grey circles, of 3 lighter and 3 darker shades relative to half grey (roughly 65, 90, 115, 140, 165, and 190 on a 255 RGB scale). The horizontal bar was now white, and the box controlled by the participant was



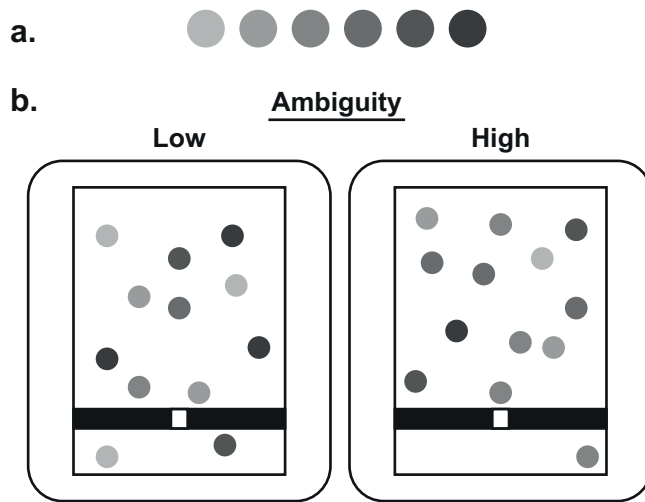
black. The number of items at the start of the trial was reduced to 6 targets and 6 distractors, in order to increase overall performance. The turbulence manipulation was adjusted to double the period of the sine wave (by dividing  $2\pi t$  by 4.8, instead of 2.4). This slowed down the rate at which the direction of the noise component added to the mouse's movement changed, thus making the mouse more controllable than in Experiment 1.

### 3.3.1.3. Design and Procedure

The main design and procedures were as in previous experiments. Instead of letters, the items consisted of grey circles. Participants were instructed to catch all the light grey circles, i.e. the targets, but to avoid the dark grey circles, i.e. the distractors (counterbalanced). Here, action selection was manipulated by varying the ambiguity in categorising items as targets or distractors. There were 6 equally spaced shades of grey, bisected by the half grey tone (see **Figure 3.4.a**). Thus, there were 3 levels of ambiguity in the grey shades, depending on the distance from half grey. The two shades at the extremes of the range were easy to categorise as light or dark, whereas the two shades closest to half grey were highly ambiguous.

Ambiguity was manipulated across trials (i.e. games), by varying the proportions of more and less ambiguous shades (see **Figure 3.4.b**). In the low ambiguity condition, most items were drawn from the extremes of the range. Thus, each trial started with 3 extreme, 2 medium, and 1 very ambiguous shade, both for targets and distractors. In the high ambiguity condition, most of the items were drawn from the highly ambiguous end of the range. Each trial started with 1 extreme, 2 medium and 3 very ambiguous shades of targets and of distractors. As before, turbulence was also manipulated across trials. As before, the 4 resulting conditions were quasi-randomised, across 6 blocks. Participants played the game for 30 s, and then gave JoAs, and JoPs.

During training, participants were shown the 3 shades of grey to class as targets, and the 3 shades to class as distractors. There were 2 practice games, both followed by JoAs and JoPs.



**Figure 3.4.** Schematic of stimuli in Experiment 2. **a.** The 6 graded shades of grey used. The first three stimuli were classed as light grey, whereas the last three were classed as dark grey. **b.** The two ambiguity conditions, which differed in their proportions of extreme and ambiguous grey shades. Stimulus size and colours have been adapted.

#### 3.3.1.4. Data Analysis

Mean  $d'$ , JoPs and JoAs were submitted to repeated measures ANOVAs with the factors ambiguity (low vs. high) and turbulence (no turbulence vs. turbulence). Simple effects t-tests were used to probe interaction between ambiguity and turbulence.

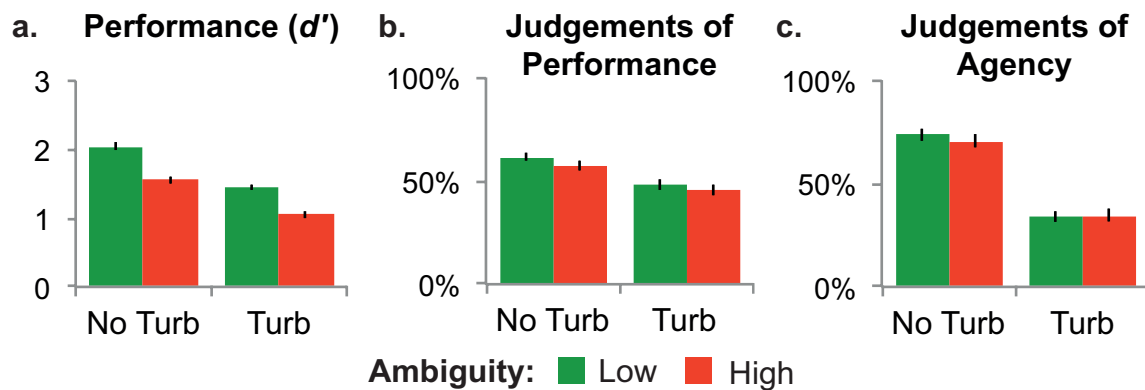
To characterise participants' strategies, the response bias ( $c$ ) measure from signal detection theory (Green & Swets, 1966) was used, with a correction for zero false alarms (Mill & O'Connor, 2014; Snodgrass & Corwin, 1988). Zero reflects an unbiased criterion, negative values reflect a liberal criterion, and positive values denote a conservative criterion. Essentially, here, a larger positive criterion would be associated with a lower number of items caught (targets or distractors), i.e. more careful responding. This bias measure was computed for each trial, and then averaged for each participant (collapsing across conditions).

Similarly to the previous experiments, JoAs were modelled by the independent variables and by JoPs. Ambiguity was coded as low = 0, high = 1. Additionally, average bias ( $c$ ) was included as a between-participant (mean centred)(i.e. level 2) covariate, as was the interaction between ambiguity and average bias.

### 3.3.2. Results

Discrimination between targets and distractors ( $d'$ ) was significantly reduced by high ambiguity, relative to low ambiguity (mean diff. = 0.45, SD = 0.20,  $F_{(1, 23)} = 116.00$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.84$ ; see **Figure 3.5.a**). Turbulence also led to significantly lower  $d'$  scores, relative to no turbulence (mean diff. = 0.55, SD = 0.14,  $F_{(1, 23)} = 362.92$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.94$ ). There was no significant interaction between the factors ( $F_{(1, 23)} = 1.57$ ,  $p = 0.22$ ,  $\eta_p^2 = 0.064$ ).

High ambiguity led to a significant reduction in JoPs, relative to low ambiguity (mean diff. = 3.04%, SD = 5.04,  $F_{(1, 23)} = 8.73$ ,  $p = 0.007$ ,  $\eta_p^2 = 0.28$ ; see **Figure 3.5.b**). Turbulence also led to significantly lower JoPs than no turbulence (mean diff. = 12.77%, SD = 8.32,  $F_{(1, 23)} = 56.62$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.71$ ). There was no significant interaction between the factors ( $F_{(1, 23)} = 1.51$ ,  $p = 0.23$ ,  $\eta_p^2 = 0.062$ ).

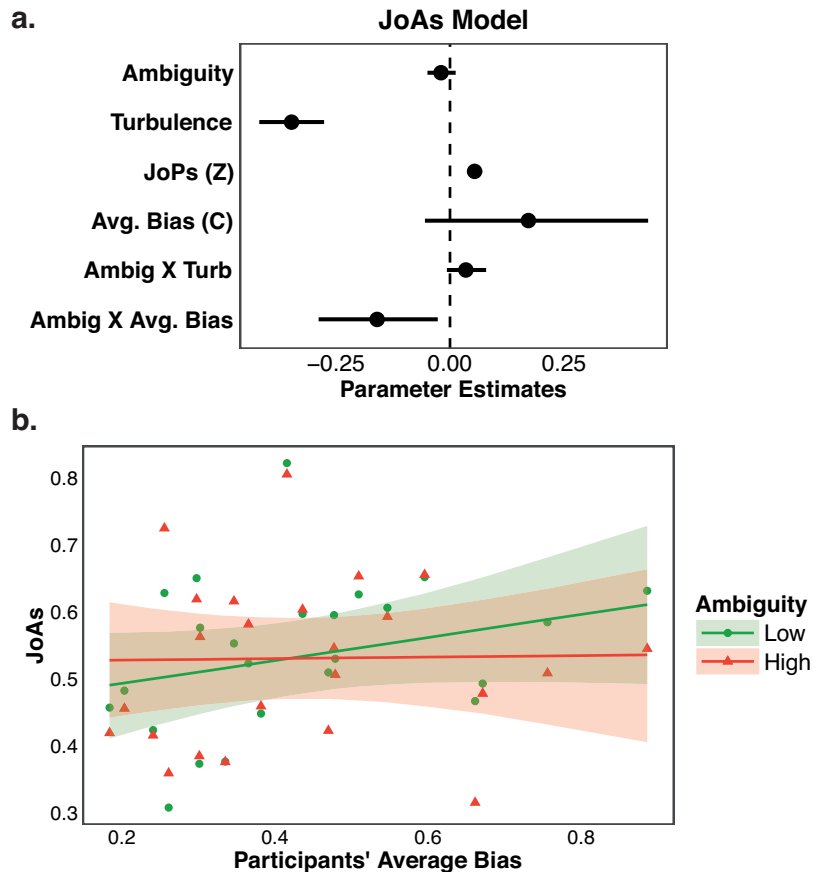


**Figure 3.5.** Results of Experiment 2. Effects of ambiguity and turbulence on mean  $d'$  (a.), JoPs (b.), and JoAs (c.). No Turb = no turbulence, Turb = turbulence. Error bars show the standard error of the mean.

Importantly, in this experiment, the effect of the ambiguity manipulation (varying the proportions of more and less ambiguous items) on JoAs could depend on the strategy participants employed in playing the game. Some might try to maximise their hits by catching many items, while risking a larger number of false alarms. Others might be very wary of false alarms, and thus focus on catching the least ambiguous items, even though they reduce their number of hits. One hypothesis would be that participants making risky choices would be more affected by the

ambiguity manipulation, because they would make more uncertain decisions, i.e. deciding to catch an item that they were uncertain about, with higher ambiguity. On the other hand, more careful participants could be more affected by the manipulation because they would be more sensitive to the overall increased uncertainty in the decision process, i.e. deciding which items to catch, in the high ambiguity condition. These strategies were captured as the average response bias (Green & Swets, 1966) of each participant, with larger (positive) values indicating more careful responding.

Therefore, in addition to assessing whether the effect of ambiguity on JoAs, could be accounted for by changes in JoPs, as in previous experiments, we assessed whether it was related to participants' average bias. For this, we added average bias, and its interaction with ambiguity, to the previously used model. As before, results showed a positive relation between JoPs and JoAs ( $b = 0.054$ ,  $t_{(22.56)} = 7.34$ ,  $p < 0.001$ , 95% CI = [0.039, 0.068]), and a significant effect of turbulence ( $b = -0.35$ ,  $t_{(23.38)} = -9.83$ ,  $p < 0.001$ , 95% CI = [-0.42, -0.28]; see **Figure 3.6.a**). Consistent with the ANOVA results, the main effect of ambiguity was not significant ( $b = -0.019$ ,  $t_{(196.38)} = -1.26$ ,  $p = 0.21$ , 95% CI = [-0.049, 0.012]). However, the ambiguity by turbulence interaction, which was significant in the ANOVA, was no longer significant in this model ( $b = 0.034$ ,  $t_{(390.44)} = 1.60$ ,  $p = 0.11$ , 95% CI = [-0.007, 0.079]). Participants' average bias was not significantly related to JoAs overall ( $b = 0.17$ ,  $t_{(22.82)} = 1.48$ ,  $p = 0.15$ , 95% CI = [-0.055, 0.43]), but, more importantly, there was a significant interaction between ambiguity and average bias ( $b = -0.16$ ,  $t_{(194.37)} = -2.60$ ,  $p = 0.010$ , 95% CI = [-0.29, -0.027]). Model predictions, displayed in **Figure 3.6.b**, showed that participants with a larger average bias, that is, with more conservative responding, showed a larger ambiguity effect.



**Figure 3.6.** JoAs Model for Experiment 2. **a.** Parameter estimates, with bootstrapped 95% confidence intervals, for modelling JoAs by the independent variables, JoPs (Z-score, within participants), and average bias (mean centred, between participants). **b.** Average JoAs across participants (points) and model predictions (regression line, and shaded 95% prediction intervals) for the relation between the effect of stimulus ambiguity on JoAs and participants' average response bias. Participants with a larger bias had a more conservative response criterion. Predictions were obtained from 10000 simulations from the posterior distribution of plausible parameter values under uniform priors (Gelman & Su, 2015).

### 3.3.3. Discussion

The present experiment investigated the effect of uncertainty in action selection on the sense of agency by manipulating stimulus ambiguity. In Experiment 1, once stimuli were identified amidst the noise mask, it was clear whether they were a target or a distractor. In contrast, items were easy to detect here, but there was uncertainty about the categorisation of highly ambiguous stimuli. Results showed that objective and subjective measures of performance were reduced by greater stimulus ambiguity, and by the turbulence manipulation. Average JoAs were also reduced by turbulence, but were only affected by stimulus ambiguity in the condition without

turbulence. That is, when the cursor accurately followed the mouse's movements, greater stimulus ambiguity led to a decrease in JoAs, but there was no difference in JoAs when the mouse and cursor movements were discrepant.

In line with the previous experiment, these findings support the proposal that metacognitive signals about action selection can influence the sense of agency (Chambon, Sidarus, et al., 2014). They further extend previous research by showing that uncertainty about the correct response to a stimulus can lead to a loss of agency. Interestingly, stimulus ambiguity reduces confidence judgements (Boldt & Yeung, 2015), suggesting there may be overlap in the signals that inform both types of metacognitive judgements.

The discrepancy between intended and observed movements, induced by mouse turbulence, again had a larger effect on JoAs than the visual manipulation (in this case, ambiguity), and, in fact, abolished stimulus ambiguity effects. A similar interaction between masking and turbulence was found in Experiment 1. The turbulence manipulation was attenuated in the present experiment, and its effect on performance was more similar in size to the effect of ambiguity, relative to the difference between the effects of turbulence and masking in Experiment 1. However, the increased overall difficulty of this task, even in the "low ambiguity" condition, may have led to weaker ambiguity effects on metacognitive judgements, as the difference between conditions was less clear.

Furthermore, this experiment allowed participants to use different strategies in playing the game. Some might risk making uncertain decisions, by catching the more ambiguous items, even though they could be distractors. Others might avoid making such risky decisions, and limit themselves to catching the less ambiguous items. To account for these differences in strategies across participants, we included a measure of the average response bias of participants when modelling JoAs. On the one hand, stimulus ambiguity might particularly affect the sense of agency when one makes more uncertain decisions. In this case, participants who made more uncertain decisions, i.e. those with a lower response bias, would presumably be more sensitive to the ambiguity manipulation. At the same time, these risky participants might have

been less concerned with whether they caught a target or distractor. On the other hand, stimulus ambiguity might have affected the sense of agency by making the task more difficult, since there were fewer easy-to-categorise items in the high ambiguity condition. Then, participants who restricted themselves to making more certain decisions, i.e. those with a higher response bias, would have been more affected by the ambiguity manipulation.

Modelling results showed that the interaction between ambiguity and turbulence found for average JoAs may be partly explained by differences in JoPs. Additionally, there was a significant interaction between ambiguity and participants' average response bias, which was independent of JoPs. This showed that stimulus ambiguity had a larger effect on JoAs in participants who had a more conservative criterion for responding, or a more positive bias, relative to less conservative participants. Participants who restricted themselves to catching those stimuli that were unambiguously identifiable as targets were most affected by the ambiguity manipulation. In the high ambiguity condition there were fewer unambiguous items, therefore the task was more difficult.

This could also reflect an overall effect of action frequency on sense of agency: conservative participants who required unambiguous evidence to identify targets for action would make relatively fewer actions during the game, especially in the high ambiguity condition. They would therefore feel a reduced sense of agency, compared to less conservative participants who made more actions. Interestingly, the interaction pattern observed in **Figure 3.6.b** shows little difference in JoAs across participants for the high ambiguity condition. Instead, it suggests that it was more conservative participants who showed an increase in JoAs in the low ambiguity condition. Perhaps these participants felt especially certain of their decisions in this condition, whereas liberal participants who made more risky decisions always felt highly uncertain about their decisions, regardless of the ambiguity condition.

These results suggest that participants' JoAs were especially sensitive to the contextual effect of stimulus ambiguity on task difficulty, i.e. whether it was easier or harder to identify targets. It remains unclear how uncertainty in a specific decision

influences the sense of agency, since the dynamic nature of the game allowed participants to avoid making more uncertain decisions. Yet, in everyday life, there are many situations in which one cannot avoid making a decision, even when one is uncertain. Thus, it could still be hypothesised that, under conditions where avoiding a decision is not possible, uncertainty about the action could influence the sense of agency.

### **3.4. Experiment 3**

This experiment tested the effect of response conflict on the sense of agency, by adding task-relevant flankers to target and distractor stimuli.

#### **3.4.1. Materials and Methods**

##### **3.4.1.1. Participants**

Procedures were as described above. Twenty-three new Columbia University or Barnard College students (18 female, mean age = 23.87, SD = 4.87) volunteered to participate for course credit, and gave written informed consent. Three participants were left-handed, 1 was ambidextrous, and the remaining were right-handed. One participant was excluded due to extremely low JoPs and JoAs ( $> 2$  SD below the mean, across conditions), and very low performance in the condition with no disruptions (full congruency and no turbulence:  $d'$  was 2 SDs below the mean).

##### **3.4.1.2. Apparatus**

The main apparatus was as in Experiment 2, except for the targets and distractors. Groups of 3 letters, consisting of Ms and Cs, were presented in white on a black background. Groups with a target letter in the central position were defined as target groups, and groups with a distractor letter in the central position were defined as distractor groups. Each trial started with 6 target groups and 6 distractor groups.

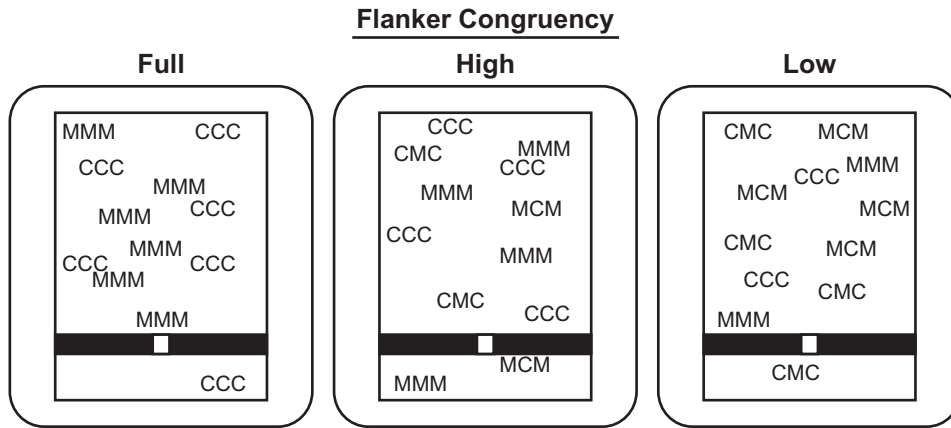


### 3.4.1.3. Design and Procedure

The design and procedure were as in previous experiments, except for the following changes. In the present experiment, Ms and Cs were presented as targets and distractors (counterbalanced). To manipulate action selection, flanker letters were added to the target and distractor items, in order to induce response conflict when flankers were incongruent with the middle letter (e.g. CMC). Items always consisted of 3-letter groups, but participants were instructed to focus on the middle letter. Only the central letter counted as a target or distractor, and participants had to touch the central letter with the mouse cursor (white box) in order to catch it. Touching only the outer letters did not count as a catch. Therefore, the flanker letters should be ignored. As before, participants played the game for 30 s, and then gave JoAs, and JoPs. Instructions were adapted for the present manipulation.

Flanker congruency was manipulated across three conditions (see **Figure 3.7**). In the full congruency condition, all items were surrounded by congruent flankers (i.e. always MMM and CCC). In the high congruency condition, two-thirds of the items were congruent, but the other third was incongruent (i.e. CMC, MCM). Finally, in the low congruency condition, only one-third of the items was congruent, and the other two-thirds were incongruent. Additionally, turbulence was manipulated across trials, as in Experiment 2. This 3 x 2 factorial design resulted in 6 conditions, which were quasi-randomised across 6 blocks, as before.

During training, participants first started by playing a game with full flanker congruency, and practiced JoAs and JoPs. Once confident with this condition, they were introduced to the incongruently flanked items, and played another practice game with high flanker congruency, followed by JoAs and JoPs. They were given the chance to practice further, if required.



**Figure 3.7.** Schematic of stimuli in Experiment 3. Stimulus size has been adapted, and colours have been inverted.

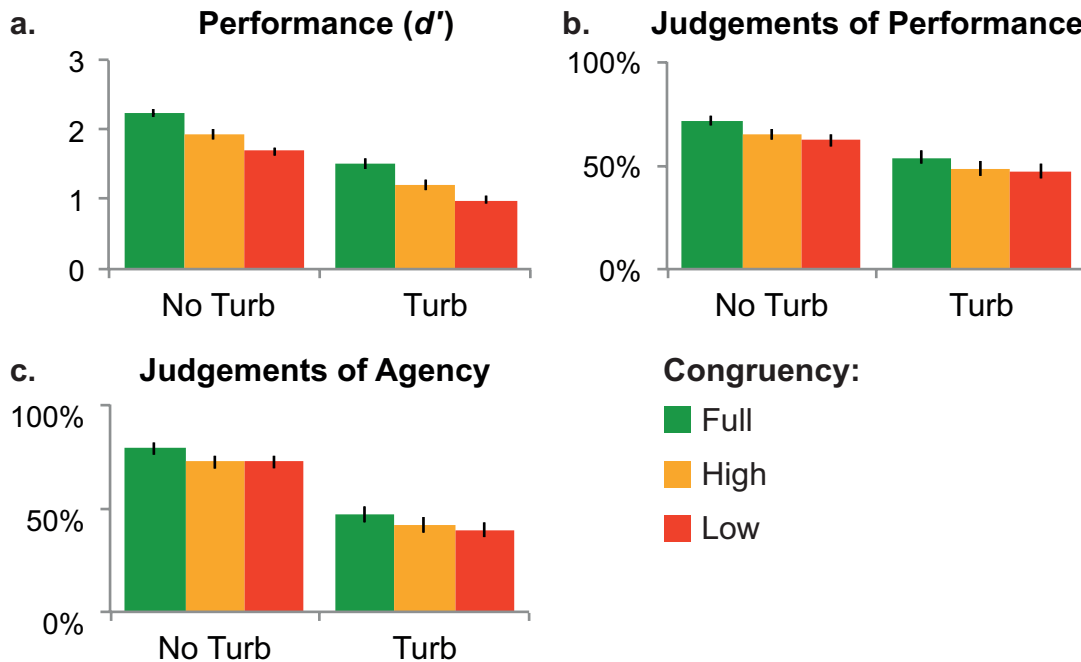
#### 3.4.1.4. Data Analysis

Mean  $d'$ , JoPs and JoAs were submitted to repeated measures ANOVAs with the factors flanker congruency (full vs. high vs. low) and turbulence (no turbulence vs. turbulence). Planned comparisons were used to probe the main effect of congruency. Additionally, JoAs were modelled by the experimental factors and by JoPs, as in Experiment 1, except for the coding of congruency. The three-level factor congruency resulted in two contrasts, with full congruency as a baseline condition (i.e. full vs. high, and full vs. low).

#### 3.4.2. Results

Discrimination between targets and distractors ( $d'$ ) was significantly affected by flanker congruency ( $F_{(2, 42)} = 105.41$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.83$ ; see **Figure 3.8.a.**). High congruency led to a significant reduction in performance, relative to full congruency (full - high: mean = 0.30, SD = 0.14), and low congruency led to a further significant reduction relative to high (high - low: mean = 0.23, SD = 0.19; all comparisons  $p < 0.001$ ). Turbulence led to a significant reduction in performance, relative to no turbulence (mean diff. = 0.72, SD = 0.21,  $F_{(1, 22)} = 253.70$   $p < 0.001$ ,  $\eta_p^2 = 0.92$ ). There was no significant interaction between the factors ( $F_{(2, 42)} = 0.23$ ,  $p = 0.80$ ,  $\eta_p^2 = 0.011$ ).

Analyses of JoPs showed a significant effect of flanker congruency ( $F_{(2, 42)} = 31.36$ ,  $p = 0.001$ ,  $\eta_p^2 = 0.60$ ; see **Figure 3.8.b.**). Relative to full congruency, high congruency led to a significant reduction in JoPs (full - high: mean = 5.87%, SD = 5.16,  $p < 0.001$ ), and there was a further reduction in JoPs for low congruency, relative to high (high - low: mean = 2.27%, SD = 4.99,  $p = 0.042$ ). Turbulence led to a significant reduction in JoPs, relative to no turbulence (mean diff. = 16.61%, SD = 11.14,  $F_{(1, 22)} = 48.49$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.70$ ). The interaction between the factors was not significant ( $F_{(2, 42)} = 0.86$ ,  $p = 0.39$ ,  $\eta_p^2 = 0.039$ , Greenhouse-Geisser corrected).

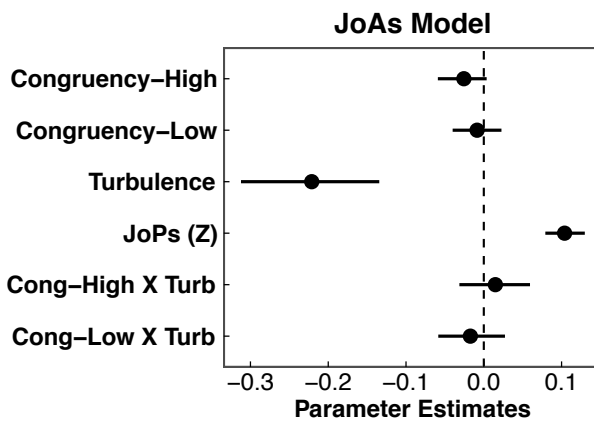


**Figure 3.8.** Results of Experiment 3. Effects of flanker congruency and turbulence on mean  $d'$  (a.), JoPs (b.), and JoAs (c.). No Turb = no turbulence, Turb = turbulence. Error bars show the standard error of the mean.

Flanker congruency also influenced JoAs significantly ( $F_{(2, 42)} = 19.91$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.49$ ; see **Figure 3.8.c.**). JoAs were significantly lower in the high and low congruency conditions, relative to full congruency (full - high: mean = 5.39%, SD = 4.80; full - low: mean = 6.66%, SD = 4.69;  $ps < 0.001$ ). JoAs did not differ between high and low congruency conditions (high - low: mean = 1.27%, SD = 6.12;  $p = 0.33$ ). Turbulence led to a reduction in JoAs, relative to no turbulence (mean

diff. = 31.48%, SD = 21.00,  $F_{(1, 22)} = 49.19$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.70$ ). The interaction between the factors was not significant ( $F_{(2, 42)} = 0.71$ ,  $p = 0.47$ ,  $\eta_p^2 = 0.033$ , Greenhouse-Geisser corrected).

Modelling JoAs with JoPs as a covariate (see **Figure 3.9**) again showed a significant positive relation between JoPs and JoAs ( $b = 0.10$ ,  $t_{(21.15)} = 7.90$ ,  $p < 0.001$ , 95% CI = [0.079, 0.13]), and turbulence remained a significant predictor of JoAs ( $b = -0.22$ ,  $t_{(20.71)} = -4.98$ ,  $p < 0.001$ , 95% CI = [-0.31, -0.13]). The contrast between full and high congruency was not a significant predictor of JoAs ( $b = -0.026$ ,  $t_{(60.60)} = -1.65$ ,  $p = 0.10$ , 95% CI = [-0.059, 0.0036]), and neither was the full vs. low congruency contrast ( $b = -0.0085$ ,  $t_{(57.39)} = -0.55$ ,  $p = 0.58$ , 95% CI = [-0.040, 0.023]). This shows that the effects of flanker congruency on JoAs could be largely explained by changes in JoPs. There were no significant interactions between the congruency contrasts and turbulence (full vs. high x turbulence:  $b = 0.015$ ,  $t_{(53.87)} = 0.69$ ,  $p = 0.49$ , 95% CI = [-0.031, 0.059]; full vs. low x turbulence:  $b = -0.017$ ,  $t_{(111.32)} = -0.83$ ,  $p = 0.41$ , 95% CI = [-0.059, 0.027]).



**Figure 3.9.** JoAs Model for Experiment 3. Parameter estimates, with bootstrapped 95% confidence intervals, for modelling JoAs by the independent variables and by JoPs (Z-score, within participants).

### 3.4.3. Discussion

In Experiment 3, action selection fluency was manipulated by varying the congruency between flankers and targets or distractors. Incongruent flankers were used to induce response conflict, as one might be mistakenly drawn towards a distractor or away from a target. Results showed that, indeed, flanker congruency affected both objective and subjective measures of performance, as well as metacognitive judgements of agency. Parametrically reducing the proportion of congruent flankers led to a gradual reduction in performance. JoPs also showed a gradual reduction in performance, but with a larger reduction when comparing full and high congruency, relative to comparing high and low congruency. This confirms that the presence of incongruent flankers led to an impairment in performance. Moreover, it suggests that, the presence of incongruent flankers was more salient than their proportion at a metacognitive level.

In fact, the results showed that average JoAs were only sensitive to the presence or absence of incongruent flankers. Relative to full congruency, both high and low congruency conditions led to a significant reduction in JoAs, but there was no difference between high and low congruency. The mere presence of incongruent flankers was sufficient to disrupt the sense of agency, independently of their proportion and, thus, independently of the precise degree of task difficulty. Interestingly, previous studies using a similar game have shown that, while JoAs are sensitive to the presence of discrepancy between the mouse's movements and visual feedback, they are not particularly sensitive to the degree of that discrepancy (Metcalf, Eich, & Castel, 2010; Metcalf, Van Snellenberg, DeRosse, Balsam, & Malhotra, 2012; Zalla, Miele, Leboyer, & Metcalf, 2015).

These results are consistent with the view that a positive sense of agency may be a "default" state, and it is when the normal flow from intention to action to outcome is disrupted that the sense of agency is reduced (Chambon, Sidarus, et al., 2014). Relative to the default state (in this case, the full congruency/no turbulence condition), any disruption is quite salient, whereas the degree of disruption may be less important. When any part of the stimulus-action-outcome processing chain is

disrupted, the sense of agency may be reduced. However, details about the disruption that informs metacognitive judgements of agency can be somewhat unspecific, not only about the locus of the disruption, but also about the degree of the disruption. Indeed, fluency/conflict signals are known to often be vague and unspecific in content, and can affect other metacognitive judgements (Winkielman et al., 2015).

As predicted, turbulence once again led to a large reduction in JoAs, which was greater than the effect of introducing incongruent flankers. Unlike Experiments 1 and 2, in which manipulations of action selection interacted with turbulence, there was no significant congruency by turbulence interaction for JoAs. As can be seen in **Figure 3.8**, there was a clear additive effect of the two manipulations on performance, and to some extent on JoPs, but the pattern was less clear for JoAs. Nevertheless, turbulence seemed to be a more important cue to agency.

Finally, modelling of JoAs revealed that they were only significantly predicted by JoPs and the turbulence factor. This suggests that changes in JoPs accounted for a large part of the congruency effects on JoAs. While this might seem to imply that flanker congruency, and thus that response conflict does not have an effect on the sense of agency independently of performance monitoring, this conclusion would be premature. Previous studies have shown that inducing response conflict with subliminal priming (Chambon, Sidarus, et al., 2014) leads to a reduction in the sense of agency independently of monitoring the external consequences of the action (e.g. the appearance of a coloured circle). Those effects were also independent of monitoring of performance in terms of RTs (Chambon & Haggard, 2012), which might be used as a proxy to selection fluency, since response conflict leads to slower RTs than fluent selection. Finally, the studies presented in Chapter 2 also used flankers to induce response conflict, and showed a reduction in the sense of agency, independent of RT and outcome monitoring.

Whereas the aforementioned studies mostly measured the sense of agency after each fluent or dysfluent action, the present experiment involved a dynamic interaction with varying proportions of congruent/incongruent flankers, and thus

assessed a more global experience of fluency vs. conflict, after the participant made several different actions in the game. The present experiment shows, as “proof of concept”, that an accumulated experience of conflict can influence the metacognition of agency. This would also be consistent with studies which obtained agency ratings at the end of a block, and showed that experiences of fluency or conflict became associated with specific action outcomes (Wenke et al., 2010; and see Chapter 2, Experiment 1).

Nonetheless, further research is needed to explore a possible dissociation between the effects of conflict on JoAs and JoPs. For example, the use of 3 congruency levels here could have led to an overall reduction in the congruency effect on JoAs, relative to an experiment comparing only full and high congruency conditions. The frequent exposure to incongruent flankers, due to the inclusion of both high and low congruency conditions, could have weakened the required stimulus-response association (i.e. catch Ms, avoid Os). This would, in turn, have weakened the conflict triggered by incongruent flankers, and thus reduced their impact on JoAs. We may therefore speculate that reinforcing the standard stimulus-response mapping by increasing the prevalence of congruent flankers could yield a greater effect of response conflict on the sense of agency.

### **3.5. General Discussion**

The present study investigated the contribution of action selection processes to the sense of agency. Over three experiments, different manipulations were used: stimulus processing fluency, stimulus ambiguity, and response conflict. Applying these manipulations in a computer game provided greater ecological validity, compared to previous studies (Chambon, Sidarus, et al., 2014). Consistently, the results showed that these various disruptions, affecting different stages of action selection, led to reductions in JoAs. These novel manipulations, applied in a dynamic environment, reveal the robustness and generalisability of the effects of action selection fluency on agency.

In addition, we probed how prospective cues, based on action selection, were integrated with retrospective cues to agency, based on proximal outcomes. For this, a discrepancy between participants' movements and visual feedback was introduced on some trials. This led to a loss of agency, in accordance with previous studies (Farrer, Bouchereau, Jeannerod, & Franck, 2008; Metcalfe et al., 2010, 2013; Metcalfe & Greene, 2007). Across all experiments, the contingency between action and (proximal) outcome, captured by the turbulence factor, remained the predominant cue to agency. Discrepant visual feedback resulted in a large cost to agency, which sometimes overshadowed the effects of action selection. When everything else was going well, and the cursor accurately followed the mouse's movements, action selection made a larger contribution to JoAs, at least when manipulating the fluency of visual processing and stimulus ambiguity. When outcomes were not as predicted, the cues provided by visual processing were less important.

In the present studies, discrepant visual feedback also resulted in a greater cost to performance and JoPs, compared to manipulations of action selection. Therefore, the similarly larger effect of discrepant visual feedback on JoAs could have been due to the fact that it caused a more salient disruption of motor control. However, in a similar game, proximal action outcomes (visual feedback of action) were also found to be a predominant cue to agency, relative to distal outcomes, such as whether the items on the screen (i.e. Xs and Os) disappeared, or not, once touched (Metcalfe et al., 2013).

These findings contrast, however, with a previous study that showed action selection fluency had a larger effect on the sense of agency when outcomes (coloured circles) violated expectations (Sidarus et al., 2013). Nevertheless, as those (distal) outcomes were mostly predictable (67% contingency), the occasional violation of expectations was not a reliable indicator of loss of agency. Under such circumstances, relying on internal cues to agency would compensate for the expected uncertainty (Yu & Dayan, 2005) of the external environment. On the other hand, here, discrepancy between the mouse and cursor movements (proximal outcomes) did reliably indicate



a loss of agency, as participants had objectively less control over the game. While the manipulations of action selection made the game more or less difficult to play, they did not affect participants' ability to interact with the game. Together, these findings suggest that sense of agency is more sensitive to proximal action outcomes (disruptions of motor control), than to distal outcomes, or action selection fluency.

Finally, as manipulations of action selection can also affect task performance, we tested whether effects on JoAs could be partially accounted for by changes in JoPs. As seen in previous studies (e.g. Metcalfe et al., 2013; Metcalfe & Greene, 2007), JoPs were positively related to JoAs. Supplementary analyses supported the assumption of a linear relation between JoPs and JoAs (see **Appendix B**), underlying the use of linear regression models. Importantly, the effects of discrepant visual feedback on JoAs could not be accounted for by changes in JoPs. Regarding manipulations of action selection, in Experiment 1, visual processing fluency had an effect on JoAs that was independent of changes in JoPs. In Experiment 2, we additionally considered the behavioural strategies of participants in playing the game, and found that the effect of stimulus ambiguity on more cautious, conservative participants was not fully accounted for by JoPs. However, in Experiment 3, the effects of response conflict on JoAs appeared to be largely explained by JoPs. Notably, the weaker effects of the action selection manipulations on JoAs in Experiments 2 and 3 may have allowed for a greater influence of JoPs

Further research is needed to clarify the interaction between prospective and retrospective cues to agency, and the role of perceived performance. Nevertheless, the present experiments present new methods for manipulating action selection that can influence the sense of agency. As action selection processes precede the outcome, they can inform the sense of agency prospectively. That is, before we even act, the process of constructing a feeling of agency has already begun. This prospective sense of agency is thought to serve as an advance predictor of successful action, and to bridge the interval between action and outcome (Chambon, Sidarus, et al., 2014). The results in this chapter confirm the role of this cue in the sense of agency, while consistently showing that the actual statistical relation

between action and (proximal) outcome remains the most important cue. A greater understanding of the prospective aspects of the sense of agency could reveal new therapeutic targets for disorders of agency, such as schizophrenia. Thus, the present study offers new avenues for investigating the prospective sense of agency.





## Chapter 4. How Action Selection Influences the Sense of Agency: an ERP study

*We investigated the neural mechanisms of prospective contributions to the sense of agency by means of event-related potentials (ERPs). Subliminal priming was used to manipulate the fluency of selecting a left or right hand action in response to a supraliminal target. Actions were followed by one of several coloured circles, after a variable delay. Participants rated their degree of control over these visual outcomes. Incompatible priming impaired action selection, and reduced sense of agency over action outcomes, relative to compatible priming. ERP components at the time of the action covaried with judgements of agency over the subsequent outcome. Feedback-related negativity evoked by the outcome was also associated with reduced agency ratings. These ERP components may reflect brain processes underlying prospective and retrospective components of sense of agency respectively. The action-related signals which prospectively influence the sense of agency are related to confidence in having selected the appropriate action.*

## 4.1. Introduction

Previous work shows that the sense of agency (SoA) is prospectively informed by a metacognitive signal about the fluency of action selection. The neural correlates of this prospective component of SoA have been studied with fMRI (Chambon et al., 2013), using subliminal priming to manipulate action selection. As this experiment was the starting point for the present work, we describe it in some detail. Participants responded to a left or right pointing arrow with a corresponding left or right hand action. After a variable delay, a coloured circle would appear. At the end of the trial, participants were asked to judge how much control they felt over that coloured circle. Importantly, and unbeknownst to the subject, a prime arrow was subliminally presented before the visible target arrow (Vorberg et al., 2003). If the prime arrow was compatible with the target, i.e. pointed in the same direction, action selection was facilitated, leading to shorter RTs. When the prime was incompatible with the target, i.e. pointed in the opposite direction, action selection was impaired, as evidenced by slower reaction times (RTs) and more errors. These incompatible priming trials were associated with lower agency ratings, relative to compatible priming. FMRI results showed that activity in the angular gyrus (AG) was related to agency ratings in incompatible priming trials, with greater AG activity being associated with a loss of agency. Notably, this activation pattern was modelled during the action selection period, between prime onset and action, thus before the outcome was known, and long before agency ratings were given.

The poor temporal resolution of fMRI does not permit a more detailed investigation of the temporal dynamics of prospective signals to SoA. In contrast to fMRI, the higher temporal resolution of EEG may help differentiate prospective processes linked to action monitoring, from later retrospective processes linked to outcome monitoring. Here we used event-related potentials (ERP) to investigate the contribution of three distinct stages of processing to SoA, locked to the target, the action, and the outcome. We next briefly consider the role of each, based on previous literature.

#### **4.1.1. Action Monitoring**

ERP studies of action priming have shown that incompatible prime-target combinations are associated with an N2 component, 200-350 ms after target onset, which is absent or greatly reduced for compatible priming (Hughes, Velmans, & De Fockert, 2009; Jiang, van Gaal, Bailey, Chen, & Zhang, 2013; Verleger & Jaskowski, 2008; Wang, Xiang, & Li, 2013). A similar component has previously been identified using the Eriksen flanker task, when distractors are incompatible with a central target (e.g. HSHH; Kopp et al., 1996). The N2 component is thought to reflect pre-response conflict detection and resolution (Donkers & van Boxtel, 2004; Yeung, Botvinick, & Cohen, 2004). N2 amplitude is linked to both the degree of conflict in a given task, and of cognitive control recruited to deal with the conflict (Larson et al., 2014). We therefore predicted that the N2 to the target stimulus might reflect action selection dysfluency, and thus be associated with prospective SoA.

A recent transcranial magnetic stimulation (TMS) study showed that disruption of inferior parietal lobe (aiming to target the AG) both before and at the time of the action, abolished action priming effects on SoA (Chambon, Moore, et al., 2014). Post-decisional action monitoring allows the integration of initial conflict signals with how well the conflict was actually handled. Crucially, action monitoring occurs in advance, and independently, of outcome monitoring. Therefore, it can provide a prospective signal to SoA, according to our definition. In action-locked ERPs, the correct related negativity (CRN) is a fronto-central component immediately following action (0-100 ms) that is thought to index post-decisional conflict monitoring, and a continued need for cognitive control (Grützmann, Riesel, Klawohn, Kathmann, & Endrass, 2014; see Larson et al., 2014 for a review). We therefore hypothesised that the Action CRN could be associated with prospective SoA.

#### **4.1.2. Outcome monitoring**

Previous ERP studies on SoA have shown that voluntary actions lead to an attenuation of outcome processing, relative to comparable externally-triggered events (Gentsch & Schütz-Bosbach, 2011; Kühn et al., 2011; Timm, SanMiguel,

Saupe, & Schröger, 2013). Sensory attenuation of outcomes has been proposed as a marker of agency also in behavioural studies (Blakemore et al., 1998; Shergill, Samson, Bays, Frith, & Wolpert, 2005). However, sensory attenuation depends on outcomes being highly predictable, close in time to the action (see Hughes et al., 2012 for a review), and high in salience (Reznik, Henkin, Levy, & Mukamel, 2015). Therefore, under conditions of uncertainty about the outcome, such as during the learning of new action-outcome associations, sensory attenuation may be less relevant to outcome monitoring.

The feedback-related negativity (FRN) is a fronto-central component seen around 250-300 ms after outcome feedback. It has been associated with reinforcement learning (Holroyd & Coles, 2002), and the learning of action-outcome associations more generally (Oliveira, McDonald, & Goodman, 2007). The FRN is typically associated with unexpected outcomes or negative feedback, such as errors or losses (San Martín, 2012). The FRN might therefore be a marker of retrospective SoA, since it reflects violations of learned action-outcome associations.

### **4.1.3. Present study**

The present study aimed to investigate these aforementioned possible neural correlates of SoA. Each trial involved subliminal action primes, supraliminal target stimuli, manual responses to those targets, a delayed visual outcome, and an explicit judgement of agency (Chambon et al., 2013). We measured candidate ERPs to different events in this sequence, to investigate potential neural correlates of prospective and retrospective components contributing to SoA. Finally, we included both free choice and instructed trials, to investigate whether the endogenous vs. exogenous origins of action contribute similarly to SoA.



## 4.2. Materials and Methods

### 4.2.1. Participants

Thirty-five participants were recruited via a UCL online database, to obtain a desired sample size of 24, based on an *a priori* power calculation (given Cohen's  $d_z = 0.65$  for within-subjects comparison of compatibility effect on agency ratings (Chambon et al., 2013), power = 0.8, alpha = 0.05). All were right-handed, with normal or corrected-to-normal vision, did not suffer from colour blindness, and had no history of psychiatric or neurological disorders. Participants received payment of £7.50/hour. Written informed consent was obtained from all participants. The study had ethical approval from the UCL Research Ethics Committee. Seven participants were excluded due to high artefact rejection rates (above 30% of the data). Three participants were excluded as they were uncooperative, or did not adequately follow instructions (e.g. repeatedly falling asleep during study; reported pre-deciding their response prior to each trial; reported in debriefing that they based their agency ratings on colour preference rather than on the relation to their own action). One further participant was excluded because they may have consciously perceived the primes (post-test  $d' = 0.65$ , over 2 SD's above the group mean  $d' = 0.02$ ,  $SD = 0.21$ ). Twenty-four participants remained (12 females, mean age = 24.38,  $SD = 4.90$ ).

### 4.2.2. Apparatus and Materials

The experiment was programmed using Psychophysics Toolbox v3 (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997). Stimuli were presented on a mid-grey background, on a 17" CRT monitor (75 Hz refresh rate) positioned at approximately 60 cm distance from participants. Prime and mask stimuli consisted of left- or right-pointing arrows, presented in dark grey. Primes subtended visual angles of  $0.8^\circ \times 1.86^\circ$ , and masks  $1.09^\circ \times 3.47^\circ$ . Prime and mask could appear randomly  $1.38^\circ$  above or below fixation to enhance the masking effect (Vorberg et al., 2003). Each action was followed by a visual stimulus in one of 8 colours (see later) subtending  $3.8^\circ$ .

### 4.2.3. Design and Procedure

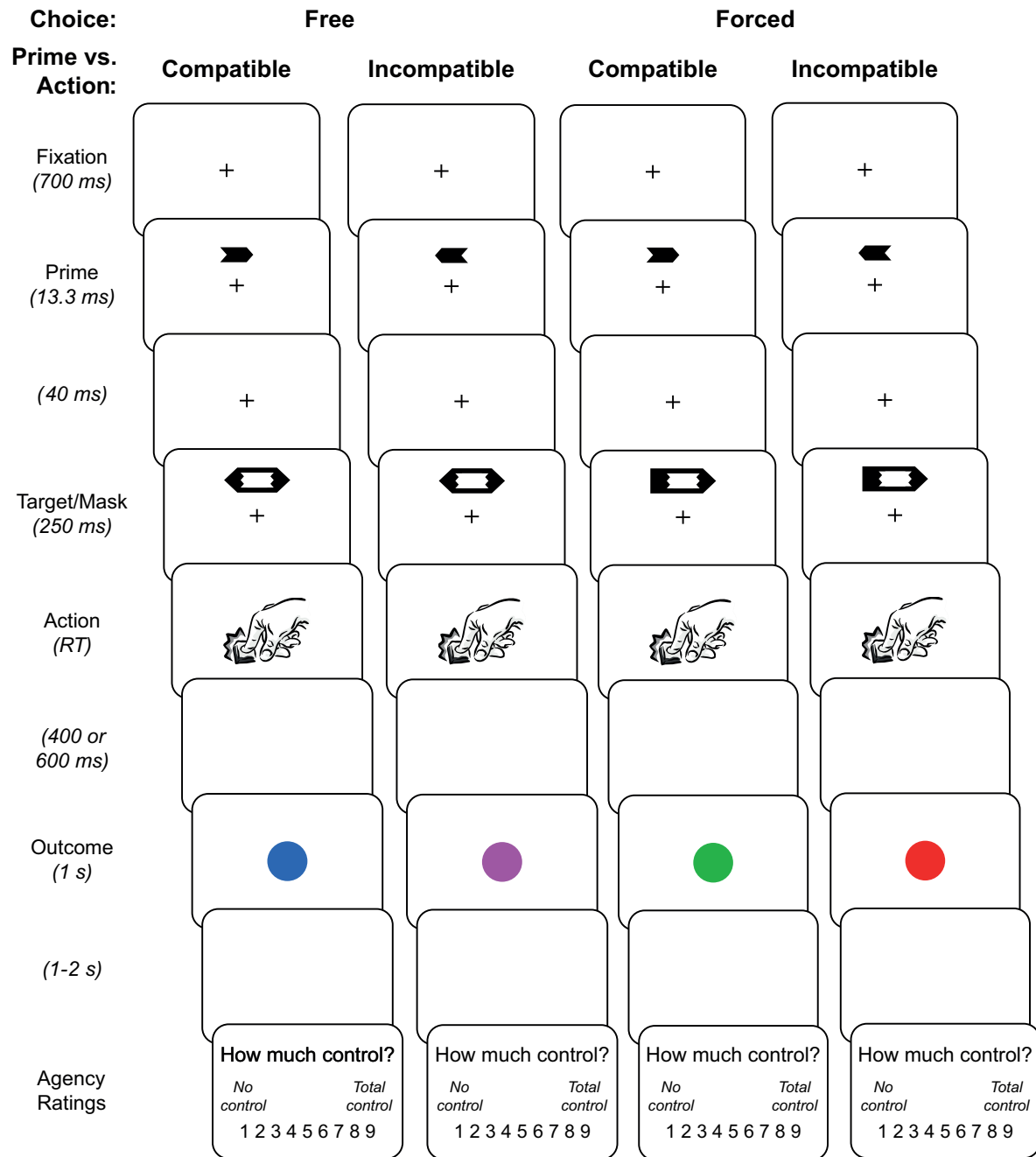
#### 4.2.3.1. Agency task

In the main task, participants had to respond to a target arrow with left or right hand action, which would trigger the appearance of a visual outcome – a coloured circle (see **Figure 4.1** below for an outline of the paradigm). At the end of a trial, they were asked to rate how much control they felt they had over the visual outcome they had just seen.

In forced choice trials, a directional arrow (pointing randomly to the left or right) instructed participants to perform the corresponding left or right hand action. In free choice trials, a bi-directional arrow indicated that participants could choose themselves whether to make a left or right hand action. Participants were instructed to try and make their choices as spontaneously as possible, and avoid deciding in advance of a trial, but at the same time, to try to choose each hand about 50% of the time. To ensure similar number of trials across actions, feedback was given at the end of each block on the percentage of left and right hand choices. Forced and free choice trials were randomised, and equiprobable.

Primes and actions could either be compatible or incompatible. In forced choice trials, prime direction could be the same as the target direction, and thus would also be compatible with the action, or primes might point in the opposite direction as the target arrow, and be incompatible with the action. Prime-action compatibility was determined online for free choice trials. When participants chose the action suggested by the prime, i.e. a left action following a left prime, trials were classed as compatible. Trials were classed as incompatible when participants chose the opposite action to the prime (i.e. a right action).

Crucially, the outcome colours were not directly related to the primes alone, or to the actions alone. Rather, the outcome colour was based on the *combination* of prime and target, or prime and action, so that compatible and incompatible trials were associated with different colours. Further, different colours were used for free and forced choices. Hence, within a block, 4 colours were associated with each choice



**Figure 4.1.** Experimental paradigm. Left or right subliminal prime arrows were briefly flashed before a target arrow, containing a metacontrast mask. Participants responded to the target by either following the instruction of directional arrows, or choosing which action to perform in response to the bi-directional (free choice) arrow. Primes and actions could be compatible (left prime – left action) or incompatible (left prime – right action). Actions triggered the onset of a visual outcome, after a variable delay. At the end of the trial, participants gave agency ratings.

condition. Of these, two colours were associated with each action, one for each level of prime-action compatibility (cf. Wenke et al., 2010). Moreover, to ensure that the frequency of each coloured outcome was equal despite differences in error rates for compatible and incompatible priming, error trials were replaced at the end of a block. Finally, to exclude any idiosyncratic preference effects, the colours were latin square rotated across 8 blocks of trials, so that each colour appeared once in each choice x action x compatibility condition. An extra block of trials with a random assignment of colours to conditions was completed at the end of the 8 blocks. This block was included to detect any anomalous use of the rating scale, but was not otherwise analysed.

The interval between action and outcome was randomly either 400 or 600 ms. The minimum interval of 400 ms between action and outcome helped to reduce the influence of action-related components on the outcome-locked ERP (Hughes & Waszak, 2011). Action-outcome interval was jittered because variation in temporal contiguity was predicted to lead to varying sense of agency, and thus to reduce stereotyped agency judgements (Haggard et al., 2002; Wenke et al., 2010). Interval duration was orthogonal to the factors of interest in the present study (choice, action and prime-action compatibility).

A trial started with a central fixation cross presented for 700 ms. Primes were shown for 13.3 ms and, after a 40 ms delay, the target/mask stimulus was displayed for 250 ms. Previous studies have shown that these parameters allow for robust priming effects, without conscious perception of primes (Vorberg et al., 2003). Participants were instructed to respond as quickly as possible to the target arrow by pressing a corresponding left or right arrow key, or to choose which action to make when bi-directional targets were shown. If they pressed the wrong key (in forced choice trials), or were too slow ( $> 1.2$  s), a black cross appeared, indicating an error. Otherwise, after a variable delay a coloured circle would appear, for 1 s. Participants were instructed to pay attention to the relation between their action and the coloured circle that followed. After a variable delay between 1 and 2 s, the rating scale was presented until participants made a response. Participants were asked to judge how

much control they felt over the coloured circle, on a Likert-type scale ranging from 1 to 9 (1 = very little control, 9 = very strong control). To prevent EEG artefacts, they were instructed not to blink until after the coloured circle disappeared. Inter-trial intervals varied randomly between 1 and 1.5 seconds.

The study started with a training block of 48 trials. If participants felt confident about the task and agency ratings, they proceeded to main experiment. The main experiment consisted of 8 blocks of 64 trials. Participants could take small breaks between the blocks. After the main experiment, participants were debriefed on the presence of primes and completed a prime awareness test.

### **4.2.3.2. Prime Awareness Test**

To assess whether primes remained subliminal for all participants, after the main experiment participants were debriefed about the presence of primes and completed a prime awareness test. This task resembled the main experiment, except without any colours following the action, or agency ratings. Participants were instructed to press the left- or right-arrow key according to the direction of the prime arrow, ignoring the supraliminal target arrow. To avoid possible response biases induced by directional targets (Vermeiren & Cleeremans, 2012), only the bi-directional arrow target was used. Additionally, a delay was introduced after mask presentation in which participants could not respond (Wenke et al., 2010). This served to prevent conscious reports from being biased by unconscious motor activations triggered by primes (Vorberg et al., 2003). This delay varied randomly between 600 and 800ms, with an auditory tone (600Hz, 150ms duration) signalling that participants could respond. This test consisted of 3 blocks of 60 trials.

### **4.2.4. EEG Recording and Analysis**

EEG was acquired with a 64 channel BioSemi Active-Two system (Biosemi Inc, Amsterdam, Netherlands) and sampled at 512 Hz. The CMS (common mode sense) and DRL (driven right leg) electrodes were used as reference and ground electrodes. Additional electrodes were placed on the left and right mastoid. Vertical and

horizontal EOGs were recorded from electrodes placed above and below the right eye and on the outer canthi of the left and right eyes.

EEG data analysis was performed with Fieldtrip (Oostenveld, Fries, Maris, & Schoffelen, 2011) and custom-built Matlab scripts (MATLAB 8.1, The MathWorks Inc., Natick, MA, 2013). All channels were 0.1-30Hz band-pass filtered, and re-referenced to average mastoids. An automatic artefact rejection procedure was employed. To identify epochs with eye-blink artefacts, EOG channels were bandpass filtered from 1-15Hz (Butterworth filter, 4<sup>th</sup> order) and any epochs with activity exceeding  $\pm 60\mu\text{V}$  were rejected. Additionally, any epochs where EEG activity exceeded  $\pm 60\mu\text{V}$  were excluded. Due to recording difficulties with some subjects, and given that these electrodes were not of interest, the following channels were excluded from analysis: T7, T8, TP7, TP8, P9 and P10. In four participants, 1 channel had to be interpolated due to abnormal noise (P7, AF3, F5 and P2 for each participant respectively). Error trials, in which participants pressed the wrong key after a forced choice target ( $M = 3.26\%$   $SD = 2.66$ ), or exceeded the response window ( $> 1.2$  s;  $M = 1.54\%$   $SD = 1.33$ ) were excluded.

**Target-locked ERPs.** Target-locked epochs were selected from 200 ms pre-stimulus to 600 ms after. Baseline correction was applied with a 100 ms interval prior to prime onset (-155 to -55 ms). Separate ERPs were calculated for each choice and priming condition (average  $N$  trials = 110, min = 64). Based on previous studies (Larson et al., 2014) and observation of grand ERPs and scalp topography, the Target N2 component was analysed as the average amplitude at Cz, between 250 and 325 ms.

**Action-locked ERPs.** Action-locked epochs were selected from 600 ms before the action to 400 ms after, with a 100 ms baseline before action (average  $N = 107$ , min = 57). Based on previous studies on the Error and Correct Related Negativity (ERN/CRN; e.g. Boldt & Yeung, 2015; Scheffers & Coles, 2000), CRN was measured as the average amplitude at FCz from 0 to 100 ms after the action.

**Outcome-Locked ERPs.** Outcome-locked epochs were selected from 200ms before stimulus to 600ms after, with a 100ms pre-stimulus baseline (average  $N$  trials = 110, min = 57). Feedback Related Negativity (FRN) was measured as the average amplitude from 250 to 300 ms at FCz, based on observation of the data and previous research (Yeung, Holroyd, & Cohen, 2005).

These ERP components were analysed with hierarchical linear regression models (also known as linear mixed-effects models). This approach, unlike classical ANOVA models, performs well with unbalanced data (Baayen, Davidson, & Bates, 2008; Bagiella, Sloan, & Heitjan, 2000; Tibon & Levy, 2015). Additionally, it allowed us to investigate the relation between agency ratings and ERP components, by modelling single-trial level data with continuous predictors. Analyses were conducted using the *lme4* package (Bates et al., 2014) in R (R Core Team, 2015). Parameter estimates ( $b$ ) and their associated  $t$ -tests ( $t$ ,  $p$ ), calculated using the Satterthwaite approximation for degrees of freedom (Kuznetsova et al., 2015), are presented to show the magnitude of the effects, with bootstrapped 95% confidence intervals. Plots of model predictions were obtained from 10000 simulations from the posterior distribution of plausible parameter values under uniform priors (Gelman & Su, 2015).

For display purposes only, but not for statistical analysis, agency ratings were median split to demonstrate the relation between neural processes indexed by ERP components and SoA. For each subject, and for each choice and priming condition, median agency rating values were obtained, and trials were classed as low or high agency ratings with respect to the median.

## 4.3. Results

### 4.3.1. Agency Task

Mean RTs were submitted to a repeated measures analysis of variance (ANOVA), with the factors choice (free vs. forced) and priming condition (compatible vs. incompatible). This revealed a significant main effect of choice ( $F_{(1,23)} = 5.92$ ,

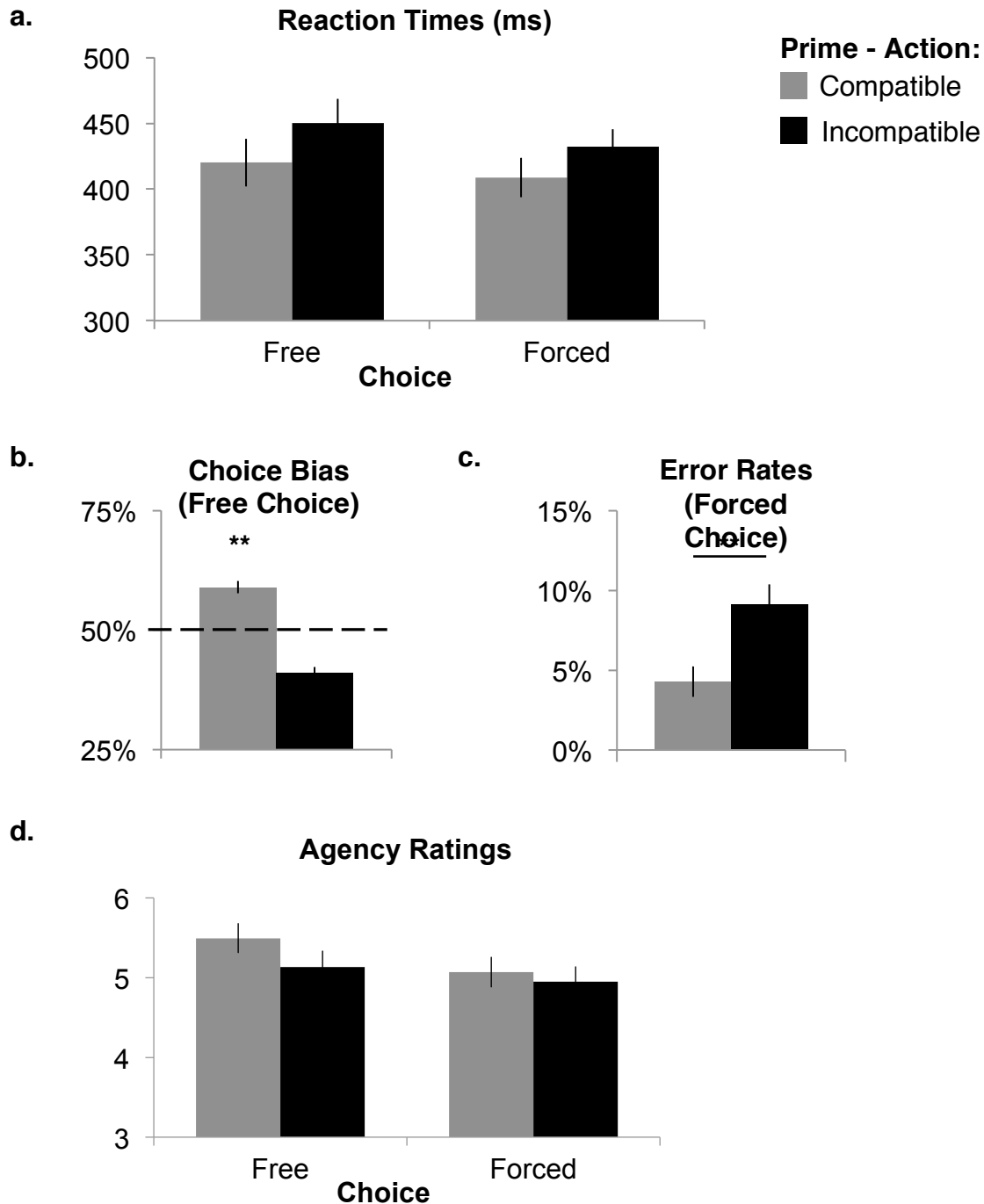
$p = 0.023$ ,  $\eta_p^2 = 0.21$ ), such that free choice trials led to slower RTs than forced choice (free:  $M = 432.11$ ,  $SD = 147.69$ ; forced:  $M = 420.65$ ,  $SD = 124.11$ ). There was also a significant main effect of prime-action compatibility ( $F_{(1, 23)} = 74.36$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.76$ ), as predicted, with slower RTs for prime incompatible actions than prime compatible actions (compatible:  $M = 414.95$ ,  $SD = 34.12$ ; incompatible:  $M = 440.52$ ,  $SD = 31.82$ ; see **Figure 4.2.a**). The interaction was not significant ( $F_{(1, 23)} = 1.98$ ,  $p = 0.17$ ,  $\eta_p^2 = 0.079$ ).

In free choice trials, prime compatible choices were made on 59% ( $SD = 0.06$ ) of trials, revealing a choice bias. A one-sample t-test revealed this was significantly different from chance level of 50% ( $t_{(23)} = 7.12$ ,  $p < 0.001$ ,  $d_z = 2.97$ ; see **Figure 4.2.b**). For forced choice trials, error rates for prime compatible and incompatible actions were submitted to a paired-samples t-test. This showed that participants made significantly more errors when primes were incompatible with the target, compared to compatible ( $M = 9.10\%$ ,  $SD = 6.26$  and  $M = 4.28\%$ ,  $SD = 4.71$ , respectively;  $t_{(23)} = -5.55$ ,  $p < 0.001$ ,  $d_z = -2.31$ ; see **Figure 4.2.c**).

Agency ratings were submitted to a 2 x 2 x 2 repeated measures ANOVA, with factors of choice (free vs. forced), prime-action compatibility (compatible vs. incompatible) and action-outcome interval (400 vs. 600 ms). Results showed a significant main effect of choice ( $F_{(1,23)} = 4.45$ ,  $p = 0.046$ ,  $\eta_p^2 = 0.16$ ), with higher ratings in free, compared to forced, choice trials (free:  $M = 5.26$ ,  $SD = 0.44$ ; forced:  $M = 5.01$ ,  $SD = 0.40$ , see **Figure 4.2.d**). A significant main effect of prime-action compatibility was also found ( $F_{(1,23)} = 10.42$ ,  $p = 0.004$ ,  $\eta_p^2 = 0.31$ ), with compatible priming leading to higher agency ratings than incompatible priming ( $M = 5.30$ ,  $SD = 0.41$  and  $M = 5.05$ ,  $SD = 0.46$  respectively). The interaction between choice and priming was not significant ( $F_{(1,23)} = 2.96$ ,  $p = 0.099$ ,  $\eta_p^2 = 0.11$ ). Finally, there was no effect of action-outcome interval, nor any interaction with the other factors ( $F_s < 1$ ). Action-outcome interval did influence agency ratings in previous studies (Chambon, Sidarus, et al., 2014), but those studies used more intervals and a wider range than the present study. Importantly, action-outcome interval was not a key



factor of interest here, and did not interact with the other factors. Therefore, it will not be discussed further.



**Figure 4.2.** Behavioural results. **a.** Mean reaction times across choice and priming conditions. **b.** Percentage of prime-compatible and prime-incompatible actions on free choice trials, showing the choice bias induced by primes. **c.** Mean error rates across priming conditions in forced choice trials. **d.** Mean agency ratings across choice and priming conditions. Error bars represent standard error of the mean. \*\* -  $p < 0.001$ .

### 4.3.2. Prime Awareness Test

Prime discrimination performance was assessed by calculating a  $d'$  score, according to signal detection theory (Green & Swets, 1966). A one-sample t-test showed that average  $d'$  did not differ significantly from zero ( $M = -0.004$ ,  $SD = 0.16$ ; two-tailed:  $t_{(23)} = -0.13$ ,  $p = 0.89$ ,  $d_z = -0.054$ ). This suggests that primes remained below the threshold of conscious awareness.

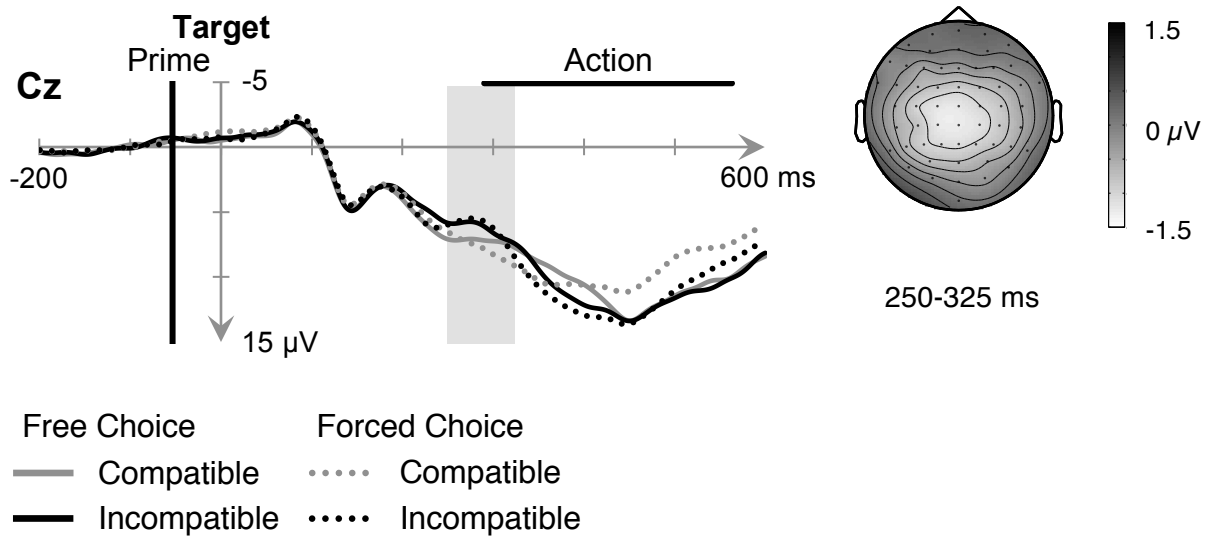
### 4.3.3. ERPs

ERP analyses focused on three neural processes that might inform SoA, at three different time points in the trial. Processes related to action monitoring were assessed at pre- and post-response stages, by analysing the target-locked N2 (Target N2) and the action-locked CRN (Action CRN) components, respectively. Finally, the outcome-locked FRN (Outcome FRN) was assessed as an index of outcome processing.

#### 4.3.3.1. Action Monitoring

##### ***Pre-Response - Target N2.***

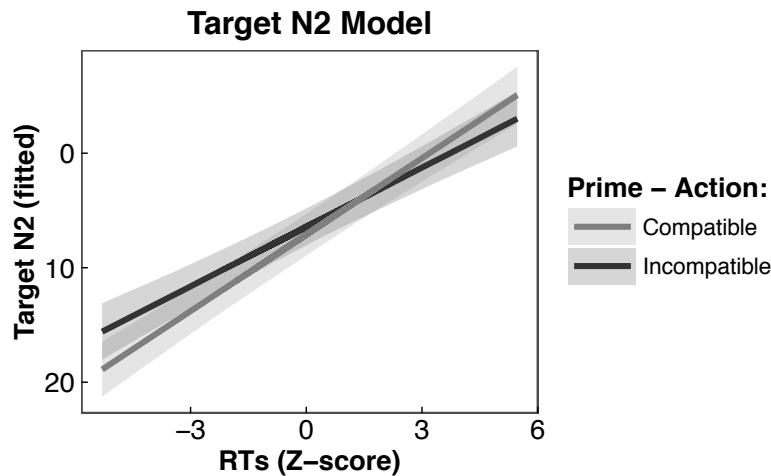
Using hierarchical linear regression, N2 amplitude was predicted from choice and priming condition (coded as 1/-1 for free/forced choice and as 1/-1 for compatible/incompatible priming), and choice by priming interaction, as fixed effects. Participants were modelled as random intercepts and random slope effects. RTs and their interactions with other factors were entered as fixed covariates, after first log-transforming the RT data, to render the distribution more normal, and standardizing it within participants.



**Figure 4.3.** Target-locked ERPs across choice and priming conditions. Incompatible priming led to an enhanced N2 component, relative to compatible priming (window highlighted in grey). This component had a central scalp distribution (incompatible – compatible subtraction). The black vertical line before the target shows the onset of the prime, and the horizontal bar indicates average RT,  $\pm 1$  SD.

The model predicting the Target N2 revealed a significant main effect of priming condition ( $b = 0.34$ ,  $t_{(34)} = 2.42$ ,  $p = 0.021$ , 95% CI = [0.080, 0.60]), a significant negative relation with RTs ( $b = -1.98$ ,  $t_{(8325)} = -17.04$ ,  $p < 0.001$ , 95% CI = [-2.18, -1.77]), and a significant interaction between priming and RTs ( $b = -0.25$ ,  $t_{(8449)} = -2.13$ ,  $p = 0.033$ , 95%CI = [-0.47, -0.0010]). No significant effect of choice, nor any other interactions were found (see **Table C1**, in **Appendix C**, for full results). These results showed that greater N2 amplitudes (more negative potentials) occurred for incompatible priming trials, relative to compatible priming (see **Figure 4.3** below for ERP and topographic plots). Additionally, greater N2 amplitudes were associated with slower RTs. To probe the interaction between priming condition and RTs, point estimates and standard errors were obtained from model predictions at  $\pm 1$  SDs of the RTs, and one sample t-tests were performed, using a conservative  $N-1$  degrees of freedom (Snijders & Bosker, 1999). The priming  $\times$  RTs interaction (**Figure 4.4**) showed that the compatibility effect on Target N2 amplitude (greater N2 for incompatible) was largest for fast RTs ( $-1$  SD RT:  $b = 1.17$ ,  $t_{(23)} = 3.17$ ,  $p = 0.004$ ), still robust at average RTs (mean RT:  $b = 0.67$ ,  $t_{(23)} = 2.41$ ,  $p = 0.024$ ), but no longer statistically significant for slow RTs ( $+1$  SD RT:  $b = 0.18$ ,  $t_{(23)} = 0.49$ ,  $p = 0.63$ ). These

results are broadly consistent with an association between the Target N2 and conflict monitoring and resolution processes



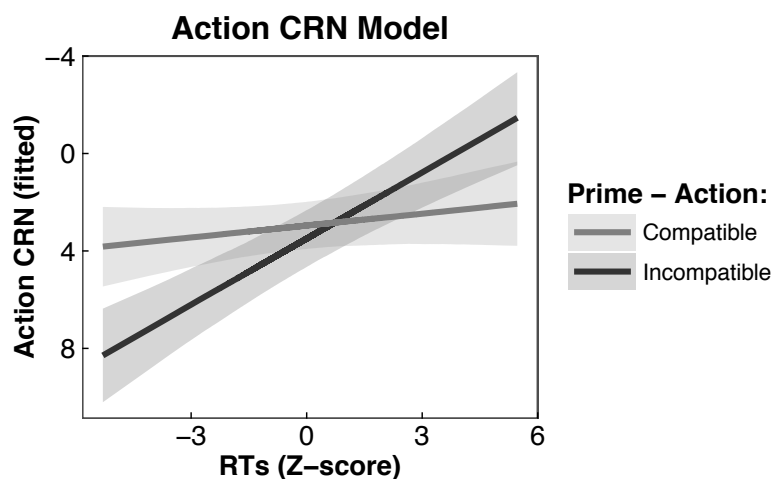
**Figure 4.4.** Target N2 model predictions for the priming by RTs interaction (with 95% prediction intervals shaded around regression lines). For fast RTs, incompatible priming led to larger Target N2 than compatible priming, but this effect reduces for slow RTs.

### ***Post-Response - Action CRN.***

The same analysis model was also applied to the mean Action CRN amplitude. The Action CRN model revealed a significant effect of priming ( $b = -0.27$ ,  $t_{(24)} = -2.12$ ,  $p = 0.044$ , 95% CI =  $[-0.51, -0.020]$ ), a significant relation with RTs ( $b = -0.54$ ,  $t_{(8485)} = -5.55$ ,  $p < 0.001$ , 95% CI =  $[-0.72, -0.35]$ ), and a significant priming by RTs interaction ( $b = .37$ ,  $t_{(8667)} = 3.85$ ,  $p < 0.001$ , 95% CI =  $[0.19, 0.55]$ ). No effects of choice, nor any other interactions were found (see **Table C2**).

Larger CRN amplitude (more negative potentials) was associated with compatible, relative to incompatible, priming. Additionally, larger CRN was associated with slower RTs. Model predictions and point estimates were again used to assess the priming by RTs interaction (see **Figure 4.5**). First, the relation between Action CRN and RTs was assessed separately for each priming condition. For compatible priming, CRN amplitude was not significantly different between fast ( $-1$  SD) and slow ( $+1$  SD) RT trials ( $b = -0.33$ ,  $t_{(23)} = -1.27$ ,  $p = 0.22$ ). However, for incompatible priming, there was a significant negative relation between RTs and CRN amplitude, such that slower RTs were associated with larger CRN than faster RTs ( $+1$  vs  $-1$  SD RT:  $b = -1.82$ ,

$t_{(23)} = -6.30, p < 0.001$ ). This revealed that the interaction between priming and RTs was driven by a modulation of CRN amplitude across RTs in incompatible priming trials, but not in compatible trials. Second, CRN amplitude was compared across priming conditions. This showed that the compatibility effect – smaller CRN for incompatible versus compatible trials – was largest for fast RTs ( $-1$  SD RT:  $b = -1.29, t_{(23)} = -3.97, p = 0.001$ ) and reduced for average RTs (mean RT:  $b = -0.54, t_{(23)} = -2.12, p = 0.045$ ). For slow RTs ( $+1$  SD), the compatibility effect was no longer significant ( $b = 0.20, t_{(23)} = 0.64, p = 0.53$ ), but reversed at very slow RTs ( $+2$  SD RT:  $b = 0.95, t_{(23)} = 2.07, p = 0.050$ ). Together, these results are consistent with an association between Action CRN and post-response conflict monitoring. They further point to the possibility that the CRN could be suppressed in incompatible priming trials in which conflict is well resolved, resulting in fast or average RTs.



**Figure 4.5.** Action CRN model predictions for the priming by RTs interaction (with 95% prediction intervals shaded around regression lines). For incompatible priming, Action CRN varied across RTs, with an enhancement of the Action CRNs for very slow RTs, but a suppression for fast RTs.

To test a possible trade-off between pre- and post-response conflict monitoring, indexed by the Target N2 and Action CRN respectively, mean Target N2 amplitude (standardised within-subjects) was added as a fixed covariate to the previous Action CRN model. Indeed, results showed a significant negative relation between Target N2 and Action CRN ( $b = -0.94, t_{(8775)} = -10.02, p < 0.001, 95\% \text{ CI} = [-1.13, -0.75]$ , see **Table C3**). Larger Target N2 (more negative potential) was associated with a smaller

Action CRN (more positive potential). Notably, the Target N2 seemed to explain some of the variance previously accounted for by the main effect of priming, as priming now became only a marginal predictor of Action CRN ( $b = -0.24$ ,  $t_{(24)} = 1.91$ ,  $p = 0.069$ , 95% CI = [-0.48, 0.011]), whereas RTs became a stronger predictor ( $b = -0.71$ ,  $t_{(8495)} = -7.27$ ,  $p < 0.001$ , 95% CI = [-0.91, -0.51]). The priming by RTs interaction remained significant ( $b = 0.35$ ,  $t_{(8663)} = -3.62$ ,  $p < 0.001$ , 95% CI = [0.16, 0.56]). These results support a distinction between pre- and post-response conflict monitoring, as the latter may integrate initial conflict signals with actual conflict resolution.

#### 4.3.3.2. Outcome Monitoring

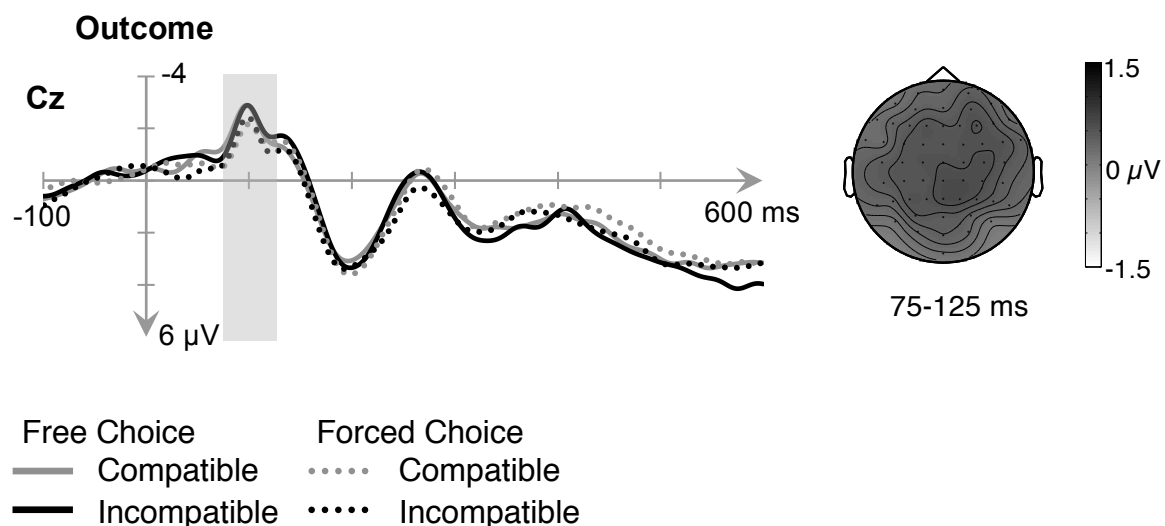
To test the hypothesis that manipulating action selection fluency may affect SoA by altering outcome processing, we modelled Outcome FRN by choice and priming condition, and their interaction, as both fixed and participant random effects. Results showed no significant effects of choice ( $b = -0.015$ ,  $t_{(22.92)} = -0.088$ ,  $p = 0.82$ , 95% CI = [-0.34, 0.34]), priming ( $b = -0.094$ ,  $t_{(42.3)} = -0.74$ ,  $p = 0.93$ , 95% CI = [-0.35, 0.16]), or choice x priming interaction ( $b = 0.13$ ,  $t_{(29.53)} = 1.00$ ,  $p = 0.33$ , 95% CI = [-0.13, 0.41]; see **Table C4**). Therefore, we found no evidence that Outcome FRN was affected by our manipulations of action selection.

Additionally, we tested a possible relation between Action CRN and Outcome FRN, by adding Action CRN to the previous Outcome FRN model. The new model revealed no significant effect of Action CRN ( $b = -0.087$ ,  $t_{(8778)} = -0.76$ ,  $p = 0.45$ , 95% CI = [-0.31, 0.12]; see **Table C5**). Since null effects cannot be clearly interpreted within frequentist statistics, Bayesian hypothesis testing (Wagenmakers, 2007) was used to further probe this relation. By comparing the Bayesian Information Criterion (BIC) between a model *with* a predictor (alternative hypothesis) and a model *without* that predictor (null hypothesis), a Bayes factor can be approximated in order to weigh the evidence *for* or *against* the null hypothesis, i.e. that the predictor is *not* related to the dependent variable. Comparing the BIC for the previous, null model with the new, alternative model yielded a Bayes factor of 70.52, indicating strong

evidence *for* the null hypothesis of no association between Action CRN and Outcome FRN.

### ***Sensory Attenuation.***

Although our task was not particularly designed to investigate sensory attenuation, we additionally tested whether our manipulations could have affected SoA by altering early sensory attenuation of outcomes. For this, we analysed the N1 component in the outcome-locked ERP (Gentsch, Kathmann, et al., 2012; Gentsch & Schütz-Bosbach, 2011). Average N1 amplitudes between 75-125 ms at Cz (Vogel & Luck, 2000) were modelled by the factors choice and priming, plus their interaction, as fixed and participant random effects. Results showed a significant negative relation between Outcome N1 and choice ( $b = -0.29$ ,  $t_{(25.30)} = -2.38$ ,  $p = 0.025$ , 95% CI =  $[-0.54, -0.033]$ , see **Figure 4.6** below), with larger (more negative) N1 amplitudes for free choice, relative to forced choice, trials. There was no effects of priming ( $b = 0.061$ ,  $t_{(269)} = 0.65$ ,  $p = 0.52$ , 95% CI =  $[-0.14, 0.25]$ ), nor choice x priming interaction ( $b = -0.0042$ ,  $t_{(25.10)} = -0.040$ ,  $p = 0.97$ , 95% CI =  $[-0.22, 0.20]$ ; **Table C6**). Therefore, we found no evidence of sensory attenuation but, in contrast, found a sensory enhancement of N1 in free choice trials.



**Figure 4.6.** Outcome-locked ERPs across choice and priming conditions. The N1 component (window highlighted in grey) was larger in free choice, relative to forced choice, trials, but was not related to priming, or agency ratings. This component had a central scalp distribution (forced – free, 75-125 ms).

As the effect of choice on agency ratings was not very robust (no longer significant after artefact rejection), it seems unlikely that this sensory enhancement is related to how action selection influences SoA. In fact, observation of **Figure 4.8.b** suggested that Outcome N1 was not related to agency ratings. Nonetheless, this was tested more directly by comparing BICs between a model predicting agency ratings by experimental factors (as above) only, with a model that additionally included Outcome N1 (standardised within-subjects). The resulting Bayes factor of 62.46 indicated strong evidence *for* the null hypothesis, that the Outcome N1 was not related to agency ratings.

#### 4.3.3.3. Predicting Agency Ratings

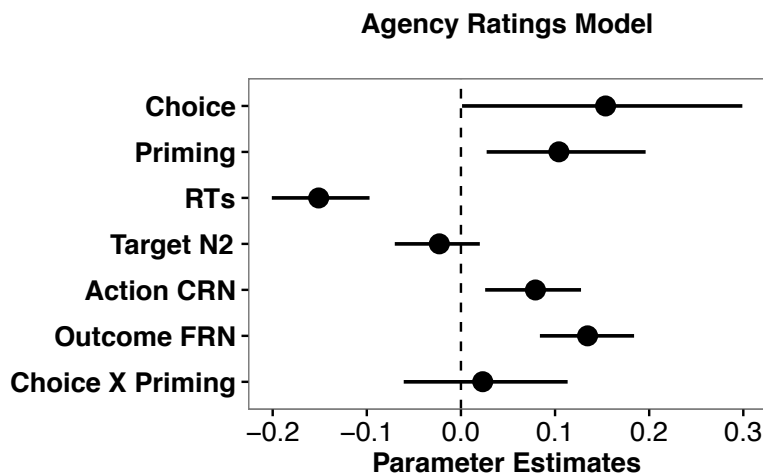
Finally, we modelled agency ratings to investigate the neural correlates of the subjective experience of agency. As above, the experimental factors of choice and priming were entered as fixed and participant random effects. Target N2, Action CRN and Outcome FRN, as well as RTs, were entered as fixed covariates (standardised within-subjects). Importantly, including the RTs as a predictor in the model ensures that other effects, such as priming effects, are estimated after taking into account the possible contribution of RTs.

Considering the experimental factors, results were consistent with the previous analysis: there was a significant effect of priming ( $b = 0.10$ ,  $t_{(23)} = 2.45$ ,  $p = 0.022$ , 95% CI = [0.021, 0.19]), a marginal effect of choice ( $b = 0.15$ ,  $t_{(23)} = 2.06$ ,  $p = 0.051$ , 95% CI = [-0.0085, 0.30]), and no significant choice x priming interaction ( $b = 0.023$ ,  $t_{(23)} = 0.56$ ,  $p = 0.58$ , 95% CI = [-0.059, 0.097]). The reduced effect of choice on agency ratings seen here, relative to the ANOVA analysis above, is due to the EEG artefact rejection procedures. Only trials with ERP data in all 3 time windows could be used for the ERP analysis, whereas all correct trials were used for behavioural data analysis.

The model further showed a significant negative relation between agency ratings and RTs ( $b = -0.15$ ,  $t_{(8726)} = -5.94$ ,  $p < 0.001$ , 95% CI = [-0.20, -0.093]), such that slower RTs were associated with lower agency ratings. This shows that RT monitoring may



partly contribute to SoA. Notwithstanding that, results show that the effect of priming on SoA could not be fully explained by differences in RTs across priming conditions.



**Figure 4.7.** Parameter estimates for the model predicting agency ratings , with 95% bootstrapped confidence intervals.

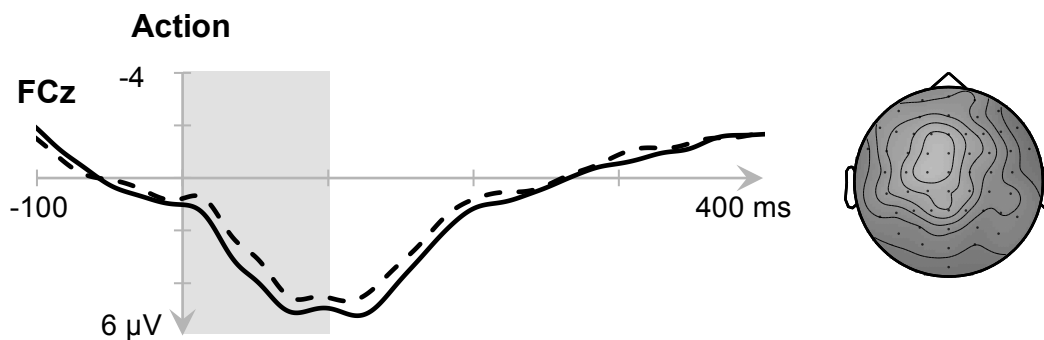
Turning to the putative neural correlates of agency, it had been hypothesised that Target N2, as a pre-response index of conflict monitoring, might be related to agency ratings. Larger Target N2 amplitudes (more negative potentials) would imply stronger response conflict and hence be associated with lower agency ratings. However, the model revealed no significant effect of Target N2 ( $b = -0.023$ ,  $t_{(8754)} = -0.92$ ,  $p = 0.36$ , 95% CI = [-0.071, 0.029]). Comparing the previous model, which included the Target N2 as a predictor, to a model without the Target N2 predictor (null hypothesis), resulted in a Bayes factor of 61.46, indicating strong evidence *for* the null hypothesis of no relation between Target N2 and agency ratings. The degree of conflict experienced, or of cognitive control recruited, during initial action selection was not directly related to SoA.

Alternatively, as an index of post-response action monitoring, it was hypothesised that Action CRN could be related to SoA. Indeed, Action CRN was found to have a significant positive relation to agency ratings ( $b = 0.079$ ,  $t_{(8760)} = 3.22$ ,  $p = 0.001$ , 95% CI = [0.029, 0.13], see **Figure 4.8.a** for ERP and topographic plots). Larger Action CRN amplitudes (more negative potentials) were associated with lower agency ratings. This would be consistent with the role of Action CRN in post-response

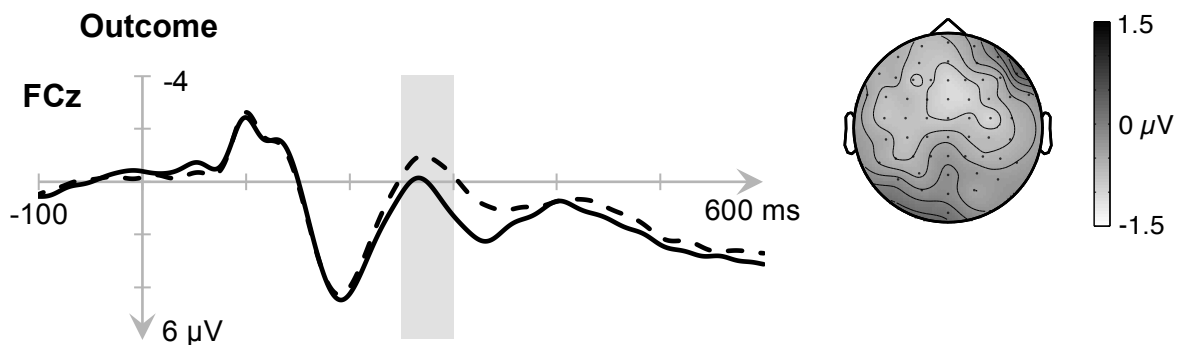
conflict monitoring, with unresolved conflict leading to greater Action CRN, and thus a reduction in SoA.

Finally, looking at neural correlate of outcome processing, we found a significant positive relation between Outcome FRN and agency ratings ( $b = 0.13$ ,  $t_{(8760)} = 5.52$ ,  $p < 0.001$ , 95% CI = [0.085, 0.18], see **Figure 4.8.b**). Larger Outcome FRN amplitudes (more negative potentials) were associated with lower agency ratings. The Outcome FRN may indicate a violation of outcome expectations, or more negative responses to the outcomes, thus leading to a reduction in SoA.

#### a. Action-Locked ERPs



#### b. Outcome-Locked ERPs



Agency Ratings

— High — — Low

**Figure 4.8.** Relation between subjective agency ratings and ERPs locked to the action and outcome onset. For display purposes, agency ratings were median split, and separate ERPs computed for high and low ratings. In panel **a**, the action-locked ERPs show that low agency ratings were associated with a larger CRN than high ratings, with a fronto-central scalp distribution (low - high agency, 0-100 ms). Panel **b**, shows the outcome-locked ERPs, revealing that low agency ratings were associated with a larger FRN than high ratings, with a fronto-central scalp distribution (low - high agency, 250-300 ms).

## 4.4. Discussion

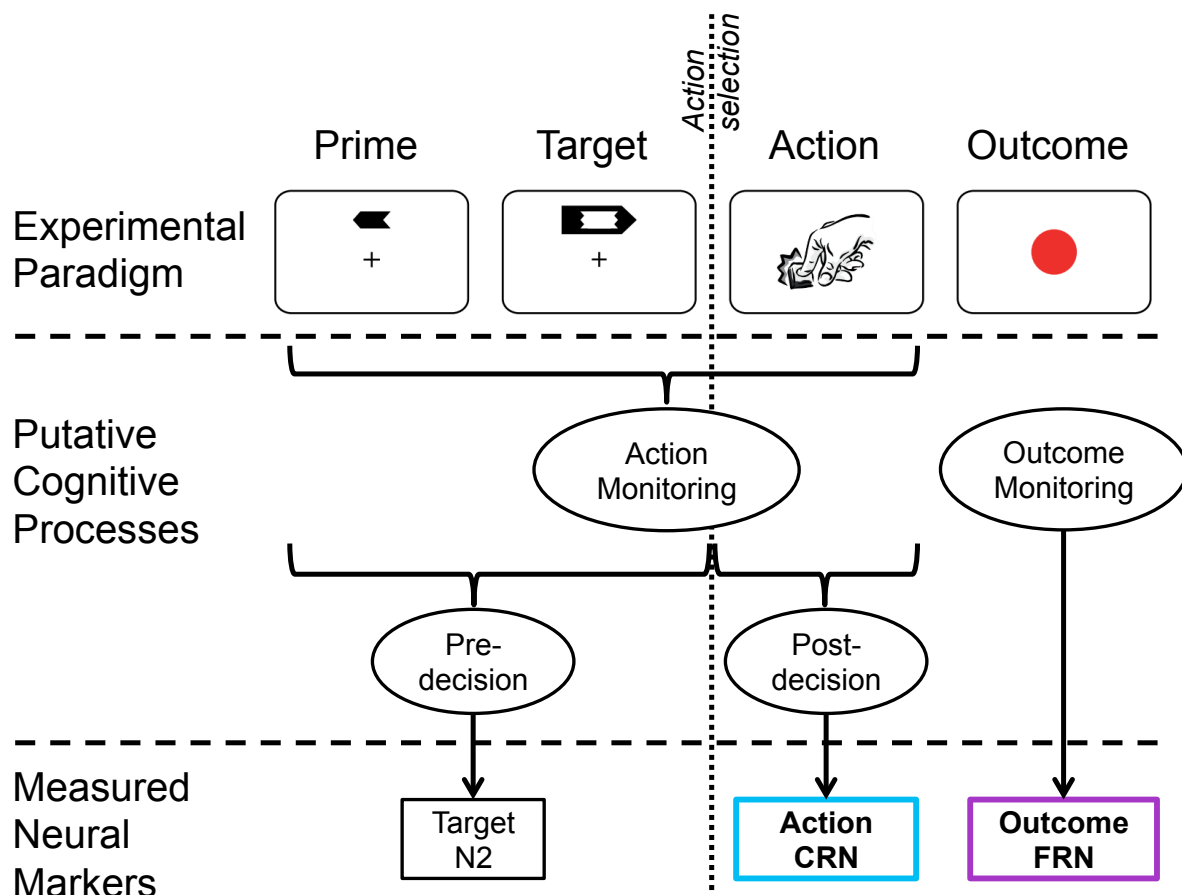
The present study aimed to clarify the neural correlates of SoA, and specifically the neural basis of prospective cues to SoA based on action selection. Behaviourally, we found that incompatible action priming led to slower RTs than compatible priming, in both free and forced choice trials. Free choices were biased towards prime-congruent choices, while incompatible priming led to more errors than compatible priming in forced choice trials. These results are consistent with previous subliminal priming studies including free choice trials (Kiesel et al., 2006; O'Connor & Neill, 2011; Schlaghecken & Eimer, 2004). More importantly, the disruption to action selection induced by incompatible priming led to a reduction in agency ratings over action outcomes, relative to compatible priming (Chambon & Haggard, 2012; Chambon, Moore, et al., 2014; Chambon et al., 2013; Sidarus et al., 2013; Wenke et al., 2010). Additionally, free choice trials led to a stronger SoA than forced choice trials, but priming had a similar effect on SoA for both choice conditions (cf. Wenke et al., 2010). At a neural level, we identified ERP components associated with SoA at the time of the action, and also at the time of the outcome. Based on previous work, these components could be identified with action monitoring and outcome monitoring, respectively (see **Figure 4.9**).

### 4.4.1. Action Monitoring

Previous studies into the neural correlates of prospective cues to SoA (Chambon, Moore, et al., 2014; Chambon et al., 2013) had not been able to disentangle the role of pre- and post-decisional action monitoring stages to SoA. The high temporal resolution of EEG allowed us to investigate this question. Broadly, components that occur early, and are related to target processing, reflect pre-decisional monitoring, while components that occur later, namely at the time of the action, reflect post-decisional monitoring processes.

Here, we found evidence of pre-response conflict monitoring, with a larger Target N2 associated with incompatible priming trials. Previous studies have shown the N2 component is sensitive to incompatibility between prime and target (Hughes et al.,

2009; Jiang et al., 2013; Verleger & Jaskowski, 2008; Wang et al., 2013), or to incompatibility between a distractor and target (e.g. Kopp et al., 1996; see Larson et al., 2014 for a review). Moreover, we found that larger Target N2 amplitude was related to slower RTs, as in previous studies (Beste, Saft, Andrich, Gold, & Falkenstein, 2008; Yeung et al., 2004; Yeung & Nieuwenhuis, 2009).



**Figure 4.9.** Schematic outline of the experimental paradigm with the putative cognitive processes involved at different stages of voluntary action. These processes are thought to be respectively indexed by the neural markers measured here. The two measures that were associated with the subjective sense of agency are highlighted. These reflect two independent cues to SoA: a prospective cue (blue outline), and a retrospective cue (purple outline).

We also found an interaction between RTs and the effect of priming on Target N2 amplitude. The compatibility effect on Target N2 (more negative amplitude for incompatible trials) was greater for fast responses than for slow responses. A previous study, that split RTs into deciles, found that the delaying effects of incompatible, relative to compatible, priming were stronger for faster RTs (Atas &

Cleeremans, 2015). Those authors proposed that trials with slower RTs may already reflect enhanced cognitive control, resulting in reduced response interference from incompatible primes. In contrast, faster responses would not benefit from the same level of cognitive control, and so would be more sensitive to the effect of the prime. Our finding of larger Target N2 associated with slower RTs for both compatible and incompatible trials strongly supports this view, and suggests that Target N2 reflects the recruitment of cognitive control. Intriguingly, the Target N2 was comparable in incompatible priming trials for both free and forced choices. To our knowledge this has not been previously investigated, but seems consistent with the observed choice bias introduced by the primes. This suggests a specific cost, or effort, is involved in ‘freely’ choosing a prime-incompatible action. Finally, we did not find any relationship between Target N2 amplitude and agency ratings, suggesting that the effects of action priming on agency were not *directly* related to the action selection stage.

We found a clear neural correlate of the SoA at the post-decisional monitoring stage. Larger Action CRN was associated with lower agency ratings. This finding is consistent with previous studies that showed a similar relation between Action CRN and metacognitive judgements of confidence (Boldt & Yeung, 2015; Scheffers & Coles, 2000). These studies showed that larger Action CRN was associated with reduced confidence in having made a correct response. Additionally, perceptual discrimination tasks that are objectively more difficult have also been associated with larger Action CRN compared to those involving easy discrimination (Endrass, Klawohn, Gruetzmänn, Ischebeck, & Kathmann, 2012; Pailing & Segalowitz, 2004). These results are consistent with a view of the Action CRN as related to post-response conflict monitoring (Grützmann, Riesel, et al., 2014; Larson et al., 2014; Yeung et al., 2004), and/or uncertainty about the correct response (Pailing & Segalowitz, 2004; Scheffers & Coles, 2000). They further suggest that prospective signals to the SoA and confidence judgements draw on information from post-decisional processes that integrate early and late action selection signals, rather than only from pre-decisional processes.

The effects of action priming on the Action CRN require further consideration. Overall, we found that compatible priming was associated with a larger Action CRN than incompatible priming. At first glance, this would seem to go against a conflict monitoring account, since that account would predict that incompatible priming trials would be associated with higher conflict, and hence larger Action CRN, than compatible priming. Yet, the literature does not support this prediction: to our knowledge, no previous studies using subliminal priming have looked at the compatibility effects on Action CRN. A few studies using the flanker task have analysed conflict effects on action-locked ERPs for correct trials, but have found inconsistent results. While one study found an overall larger Action CRN for compatible, relative to incompatible, flanker trials (Grützmann, Riesel, et al., 2014), others have shown non-significant effects in the same direction (Cohen & Donner, 2013; Scheffers & Coles, 2000), while yet another study has reported the opposite effect (larger Action CRN for incompatible; Bartholow et al., 2005). In line with a conflict monitoring account, our results also showed that a larger Action CRN was associated with slower RTs. Importantly, there was a significant interaction between priming and RTs, qualifying the two main effects. The relationship between Action CRN and RTs was specifically present for incompatible priming trials, whereas Action CRN was stable across fast and slow RTs for compatible priming trials.

When priming is incompatible, a fast RT might reflect a trial in which the prime was less processed, leading to a reduced, or absent, response conflict. However, this would predict no difference in Action CRN between compatible and incompatible trials. In fact, the current results show that incompatible priming was associated with smaller Action CRN than compatible priming for fast and average RTs. Alternatively, incompatible priming trials with fast RTs could reflect trials with faster and/or more efficient recruitment of cognitive control processes to overcome response conflict. Then, the observation of a smaller Action CRN for incompatible priming in trials with fast RTs could be interpreted as a suppression of Action CRN due to efficient cognitive control deployment before the action. This suppression would not be necessary in compatible priming trials. While our findings cannot disambiguate between these hypotheses, they do point to a complementarity, or trade-off, between

pre- and post-response conflict monitoring (cf. Grützmann et al., 2014; Larson et al., 2014).

Both the pre-response N2 and post-response negativities (Action CRN/ERN) have been linked to the anterior cingulate cortex (ACC) (van Veen & Carter, 2002; Yeung et al., 2004), suggesting they reflect two successive recruitment events of a single conflict monitoring circuit. If conflict is adequately detected and resolved before the response, there may be no need for further conflict-related processing after the response. However, if conflict is not fully processed before the response, post-response conflict signals could be important to help prevent future errors. To test this complementarity in our data, we added Target N2 amplitude as a covariate to our model predicting Action CRN amplitude. Results indeed showed that larger Target N2 was associated with smaller Action CRN, consistent with a previous study (Grützmann, Riesel, et al., 2014) that showed this relation both at the within- and the between-subject level. Together, these results support the view that post-response action monitoring can integrate initial conflict signals with actual conflict resolution. The Action CRN may reflect a persisting need for cognitive control recruitment at the time of the action (Grützmann et al., 2014). Yet, successful cognitive control at an early, pre-decisional stage would reduce or obviate the need for later, post-decisional cognitive control.

Prospective SoA may therefore be linked to a post-decisional integration of signals both early and late in the action-generation process (see **Figure 4.9**). A putative role of the prospective component of SoA may be as an experiential marker, or an epistemic feeling (Proust, 2008), of the unfolding voluntary action. If action selection is dysfluent, there may be a need to adjust behaviour. If an error is made, corrective action may help to avoid unintended outcomes. Even when the correct action is made, a signal that “something went wrong” (Pacherie, 2008), might weaken the link between the action and ensuing outcome, leading to a reduced SoA over the outcome. When processing action-outcome contingencies, it may be adaptive to *learn* following fluent, high-confidence actions, but *not learn* following dysfluent, low-confidence, actions.

Alternatively, conflicts can serve as aversive signals (Botvinick, 2007). Conflict induces negative evaluations of subsequent neutral stimuli (Fritz & Dreisbach, 2013), and triggers behavioural adjustments in subsequent trials (Dreisbach & Fischer, 2011, 2012). In fact, post-response negativities (CRN/ERN) have been linked to negative affect (Hajcak, McDonald, & Simons, 2004; Simon-Thomas & Knight, 2005), and to the motivational significance of errors (Aarts, De Houwer, & Pourtois, 2013; Grützmann, Endrass, Klawohn, & Kathmann, 2014; Hajcak & Foti, 2008). From this perspective, our finding, of reduced SoA over outcomes following response conflict, could be interpreted as the negative valence of conflict “leaking” into outcome evaluation.

### **4.4.2. Outcome monitoring**

After outcome onset, we found that larger (more negative) Outcome FRN amplitudes were associated with lower agency ratings. Previous studies comparing Outcome FRN between gains and losses, have shown greater FRN sensitivity associated with outcomes contingent on action vs. non-contingent (Yeung et al., 2005); and with greater perceived responsibility over outcomes (Li et al., 2010; Li, Han, Lei, Holroyd, & Li, 2011). Our results are also consistent with a recent report that stronger negative potentials in a similar time window were associated with outcomes that were externally- vs self-attributed (Bednark & Franz, 2014). More generally, the terminology for this component has been disputed. On one view, this component would be best described as a positive-going potential for positive outcomes, termed feedback-correct related positivity (fCRP; Holroyd, Pakzad-Vaezi, & Krigolson, 2008; Oliveira, McDonald, & Goodman, 2007; see San Martín, 2012 for a review). Irrespective of naming conventions, our results, and those discussed above, agree that less negative (or more positive) potentials for correct or rewarding outcomes are a putative correlate of greater SoA.

Previous EEG studies on agency have manipulated outcome expectation or predictability, for example, by priming the outcome (Gentsch, Kathmann, et al., 2012; Gentsch & Schütz-Bosbach, 2011), or by violating learned action-outcome



contingencies (Kühn et al., 2011). These studies have found sensory attenuation to self-triggered and predicted/expected outcomes in early (N1) processing; as well as smaller P3a for self-attributed, relative to externally-attributed, outcomes (Kühn et al., 2011). Additionally, using subliminal action priming with fully predictable outcomes, a study has found a sensory attenuation of outcomes that followed compatibly primed, relative to incompatibly primed, actions (Stenner, Bauer, Sidarus, et al., 2014).

In contrast, we manipulated action selection while equating outcome predictability, and frequency, across conditions (choice x priming). Here, outcomes were always relatively uncertain, as there were many possible outcomes (8), and action-outcome contingencies changed across blocks. Within a block, participants could learn that left and right hand actions were each associated with 4 colours, 2 for free and 2 for forced choice trials. The compatibility relation between prime and target disambiguated the remaining 2 colours, but this information was not available since the primes were subliminal. Sensory attenuation depends on the ability to adequately predict a sensory event (see Hughes et al., 2012 for a review), therefore such measures were not particularly suited to our design.

Our analyses did not show any evidence of sensory attenuation associated with the priming manipulation. We did find that free choices were associated with higher N1 amplitude, relative to forced choices. The N1 component is well known to be enhanced by attention (Vogel & Luck, 2000), suggesting that freely choosing what to do may enhance attention to action consequences. Consistently, a recent study found larger N1 for the auditory outcomes of free choices, relative to a coercive condition (Caspar et al., 2016). Importantly, in the present study, the effect of choice on agency ratings was not very robust, and became only marginally significant after EEG artefact rejection. In contrast, the effect of priming was highly robust. Moreover, we found no association between Outcome N1 amplitude and agency ratings, as can be observed in **Figure 4.8.b** of the main article.

Despite considerable outcome uncertainty, participants could still learn action-outcome associations throughout a block. This may be reflected in the present Outcome FRN findings. The expectancy-deviation account of the FRN (Oliveira et

al., 2007) proposes that the FRN is associated with ACC-mediated monitoring and learning of action-outcome associations. Mismatches between expected and observed outcomes would result in greater ACC activation, and larger FRN, signalling a need for cognitive control and an updating of internal models of action-outcome contingencies. Similarly, the reinforcement learning account (Holroyd & Coles, 2002; Holroyd et al., 2008) suggests that the FRN reflects dopaminergic prediction-error signals. Negative prediction-errors, or worse than expected outcomes, would lead to more negative FRN potentials. Therefore, our findings could be due to a larger Outcome FRN being evoked in trials with outcome prediction errors, which were in turn related to lower agency ratings. Our results further showed no effects of priming or choice on Outcome FRN. This suggests that priming did not affect agency by directly altering outcome processing, namely by disrupting outcome predictions.

The FRN has been widely linked to negative outcomes (San Martín, 2012), such as monetary losses (Gehring & Willoughby, 2002), or no reward relative to reward (e.g. Holroyd, Krigolson, Baker, Lee, & Gibson, 2009). Therefore, the present Outcome FRN findings could also reflect varying affective evaluations of the outcome. As mentioned before, response conflict can be considered aversive, and influence affective responses to ensuing events (Fritz & Dreisbach, 2013). Disruptions in action selection, induced by incompatible priming, could have led to more negative affective responses to action outcomes, leading to larger Outcome FRN and lower SoA.

Importantly, the relation between Outcome FRN and SoA was independent of a possible link between Action CRN and conflict-induced negative affect. In modelling of agency ratings, we found no relation between these two ERPs, suggesting that the Outcome FRN did not *directly* reflect an affective response linked to the preceding action. Furthermore, Action CRN and Outcome FRN were significant predictors of agency ratings, suggesting that they explained different portions of the variance. Outcome monitoring, as indexed by the Outcome FRN, may be partly influenced by action monitoring signals, but it also integrates information about the observed

outcome, and whether it matches internal models of action-outcome contingencies. Our study supports the proposal that prospective, action selection-based, and retrospective, outcome-based, cues can make independent contributions to the SoA (Sidarus et al., 2013), at least when outcomes are not highly predictable (Stenner, Bauer, Sidarus, et al., 2014; Stenner, Bauer, Heinze, Haggard, & Dolan, 2014).

### **4.4.3. Common performance monitoring framework**

Converging evidence shows that response-related negativities, i.e. Action CRN/ERN, and the Outcome FRN have common neural mechanisms, linked to the ACC (Botvinick, 2007; Frank, Woroch, & Curran, 2005; Holroyd & Coles, 2002; Larson et al., 2014; San Martín, 2012). The ACC is thought to be involved in goal-directed action, driving action-outcome learning and adaptive behaviour (Botvinick, 2007; Holroyd & Yeung, 2012). Monitoring actions and outcomes for response conflict, errors, and negative or unexpected outcomes, the ACC can signal a need for cognitive control. Structures such as the dorsolateral pre-frontal cortex can in turn adjust current behavioural strategies or internal models of action-outcome associations.

Notwithstanding this commonality, action and outcome monitoring use different information, available at different times. Action monitoring relies on internal, prospective signals about action selection and execution, whereas outcome monitoring depends on external, retrospective feedback processing in sensory areas. Here we saw that the SoA was independently related to both Action CRN and Outcome FRN, consistent with the idea that they integrate different information. Interestingly, a study has reported a disruption in ERN but intact FRN in patients with schizophrenia (Horan, Foti, Hajcak, Wynn, & Green, 2012), thus dissociating these two monitoring processes, and potentially their role in SoA. Other studies have similarly suggested that schizophrenia patients have impaired monitoring of internal, action-related signals, and are instead over-reliant on external, outcome monitoring (Metcalfe, Van Snellenberg, et al., 2012; Voss et al., 2010). Furthermore, the ERN is sensitive to whether errors are internally or externally generated, e.g. a response

button malfunction (Gentsch, Ullsperger, & Ullsperger, 2009; Padrao, Gonzalez-Franco, Sanchez-Vives, Slater, & Rodriguez-Fornells, 2016; Steinhauser & Kiesel, 2011). Internally-generated errors led to a large ERN, whereas externally-generated errors were associated with later ERP components, arguably FRN-like (Gentsch et al., 2009).

Interestingly, with the exception of studies on error monitoring, the ACC and ACC-mediated performance monitoring have rarely been linked to SoA. The agency literature has typically employed considerably different tasks from the one used here, focusing especially on the *attributional* aspect of agency, i.e. “who did it” (Chambon, Filevich, et al., 2014). For example, participants may be asked to judge whether outcomes were caused by oneself or another agent. This research has linked sensorimotor control and outcome monitoring to the parietal cortex, though other frontal and premotor areas have been implicated (David et al., 2008; Farrer, Frey, et al., 2008; Fink et al., 1999; Miele et al., 2011).

In contrast, our study focused on the *instrumental* aspect of agency, involving monitoring and using more abstract action-outcome associations. As the ACC is involved in goal-directed actions and performance monitoring more generally, it is especially relevant to this aspect of SoA, rather than to agency attribution. The SoA is complex and multifaceted, involving the integration of multiple signals, from internal sensorimotor signals, to external feedback, to higher-level beliefs and inferences (Gallagher, 2012; Synofzik, Vosgerau, & Voss, 2013). We speculate that regions previously associated with SoA, may integrate signals from the ACC with other inputs relevant to determining agency. The AG is a likely candidate for this integration as it has been linked to a subjective loss of agency associated with dysfluent action selection (Chambon et al., 2013), as well as with unexpected outcomes (Farrer et al., 2003; Farrer, Frey, et al., 2008; Farrer & Frith, 2002; Nahab et al., 2011). The AG may subserve an online monitoring system that tracks the whole intentional action chain, from intentions, to actions, to outcomes, and signals a loss of agency.

#### 4.4.4. Conclusions

We found that an unconscious influence on action selection processes, from subliminal priming, can affect the conscious experience of agency over action outcomes. ERP results showed that action monitoring signals influence SoA prospectively, since the neural correlates of SoA emerge at the time of the action, long before the outcome is known. The association seen here between Action CRN and agency ratings mirrors associations found between Action CRN/ERN and confidence ratings (Boldt & Yeung, 2015; Scheffers & Coles, 2000). This suggests that the signals related to action selection which influence SoA could be better described as relating to *confidence* in selecting or having selected the appropriate response, and not only to *selection fluency* as has been previously described (Chambon, Sidarus, et al., 2014; Wenke et al., 2010). Our results therefore link prospective sense of agency to the processes of action monitoring and cognitive control. These results invite speculations, about possible functions of SoA within human cognition generally. In particular, the SoA may provide an important experiential marker, both for alerting to the need for corrective action, and for guiding learning.



## **Chapter 5. Integrating Prospective and Retrospective Cues to the Sense of Agency: A Multi-Study Investigation**

*Sense of agency (SoA) involves a complex integration of various cues. These include prospective, related to the fluency of action selection, and retrospective, linked to outcome monitoring. It remains unclear whether these cues may have independent effects on SoA, and, in particular, how their relative contributions may change during instrumental learning. In the present study, we investigated these issues by conducting a multi-study analysis, combining the experiments described here in Chapters 2 & 4, with other known studies on prospective cues to the SoA (with a total of seven experiments).*

*Our main question was how the effects of selection fluency on SoA might change as information about action-outcome contingencies is gathered. Results show that selection fluency can have a general, and consistent influence on the SoA, independent of outcome monitoring. Selection fluency is used as a heuristic cue, to prospectively inform our SoA. In addition, our results show that the influence of selection fluency on SoA may change during the learning of action-outcome contingencies. We speculate that dysfluent selection may impair the linkage between action and outcome.*

## 5.1. Introduction

The sense of agency (SoA) is highly sensitive to disruptions in the chain of events of voluntary action (Chambon, Sidarus, et al., 2014; Gallagher, 2012): SoA emerges from establishing a link between our intentions, actions, and external outcomes. Importantly, the SoA involves the integration of multiple signals, which may become available at different stages of voluntary action (Chambon, Sidarus, et al., 2014; Farrer et al., 2013; Moore & Fletcher, 2012; Synofzik et al., 2013). Monitoring the fluency of action selection processes can provide an initial, prospective cue to SoA. Forward model predictions about the outcome can then be compared with the observed outcome, to retrospectively link action and outcome. These predictions may concern both the interval between action and outcome, and outcome identity. Additionally, higher-level beliefs and contextual information can also influence the SoA.

Some have proposed that this integration follows the principles of optimal cue integration (Moore & Fletcher, 2012; Synofzik et al., 2013): cues are weighted based on their reliability, thus more reliable cues have a stronger influence than less reliable cues. Additionally, the weighting of cues may be altered by prior knowledge, or contextual cues. This approach has previously proved useful to understanding cue integration in SoA (Moore & Haggard, 2008; Wolpe et al., 2013). Yet, how the weighting of different cues may change dynamically with experience, for example, throughout the process of learning new action-outcome contingencies, remains poorly understood.

It has also been suggested that selection fluency may become a useful cue to SoA because it is predictive of successful action (Chambon, Sidarus, et al., 2014; Haggard & Chambon, 2012). Yet, it remains unclear how the process of learning about action-outcome relations may influence the role of prospective cues to SoA. On the one hand, the relation between selection fluency and particular outcomes might need to be acquired in specific contexts, in order to inform SoA. On the other hand, lifelong experience might establish a link between selection fluency and



successful action, and thus selection fluency could be generally used as a heuristic cue to SoA.

Alternatively, the fluency of action selection might itself influence the learning of action-outcome contingencies. Learning to associate dysfluent actions to their outcomes could be impaired, relative to learning about the outcomes of fluent actions. In line with this perspective, sensorimotor predictions of action consequences can be disrupted by incongruent subliminal priming of actions (Stenner, Bauer, Sidarus, et al., 2014; Stenner, Bauer, Heinze, et al., 2014).

In the present study, we investigated the relation between prospective and retrospective cues to SoA, based on action selection and outcome monitoring, respectively. More specifically, we assessed how the contribution of selection fluency to SoA may change during instrumental learning, as information about action-outcome contingencies is gathered. For this, we conducted a multi-study analysis, combining the experiments described here in Chapters 2 & 4, with currently known studies on prospective cues to the SoA.

In these studies, three cues to SoA were varied (see **Figure 5.1.A**). Selection fluency was manipulated by varying the congruency between primes, or flankers, and the executed action. The action was followed by a variable action-outcome interval (AOI). The action outcome (one of several coloured circles) depended on the action and congruency conditions. Therefore, in each trial, these three cues could be combined to inform agency ratings (our measure of SoA). Since action-outcome relations had to be learned anew in each block of trials, tracking changes in agency ratings across trials indexed the contribution of monitoring outcome identity.

Importantly, the relative contribution of these 3 cues to SoA could be modulated by contextual information, such as instructions. If participants are instructed to focus on a particular cue to SoA, e.g. outcome identity, the contribution of that cue to SoA would likely increase overall. Additionally, this could also alter the contribution of the other cues to SoA. Within the experiments analysed here, there were two groups of studies that differed in the instructions given to participants about the agency ratings

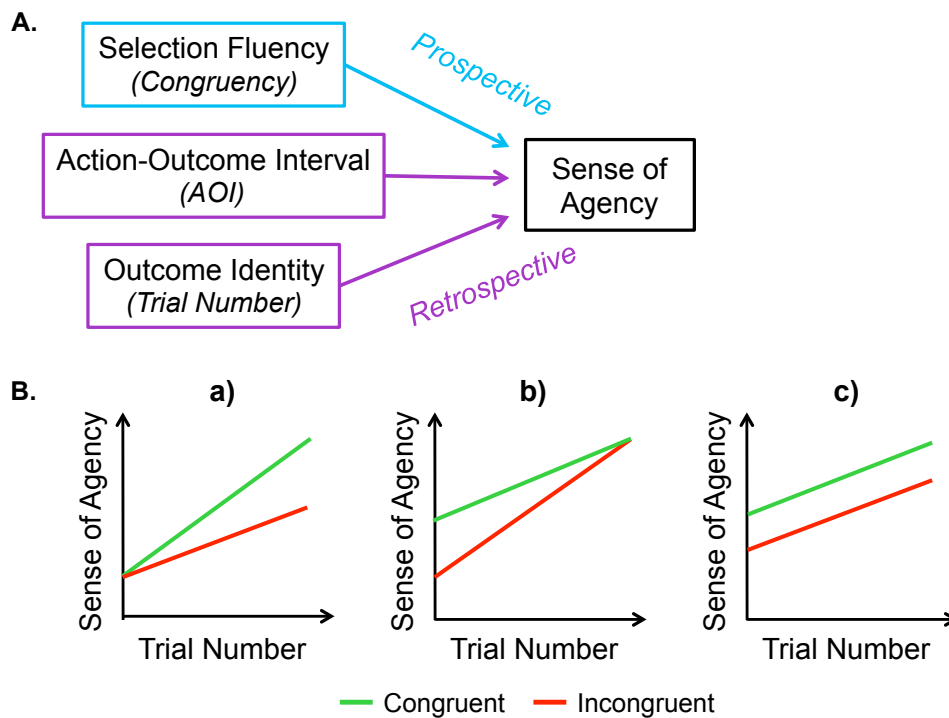
procedure. The studies in Group 1 (in Chapters 2 & 4) instructed participants to focus on the relation between actions and outcomes, that is, to focus on outcome identity. Studies in Group 2 additionally instructed participants to consider AOI (see **Table 5.1** in Methods).

In general, assuming participants are attending to action-outcome relations, as instructed, we predict that agency ratings would increase across trials. If that is the case, we can consider possible interactions between selection fluency and outcome identity. Given the literature discussed above, our main question was whether we must learn to use selection fluency as a cue to SoA in specific contexts, or whether it may be a heuristic learned from everyday experience. We suggest 4 different accounts on the relation between selection fluency and outcome identity, which could be consistent with three possible results for how the effect of selection fluency on agency ratings may change across trials (see **Figure 5.1.B**).

1. *Learning to be prospective.* As we learn specific action-outcome relations in each block, we become able to use selection fluency as a proxy for causing a specific outcome. This predicts that the effects of selection fluency on SoA would increase across trials (pattern **a**).
2. *Optimally prospective.* We rely on the heuristic of selection fluency as a cue to SoA to prospectively guide our SoA at the start of a block, when outcomes are not a reliable cue. Once enough knowledge is gathered about action-outcome contingencies, outcomes will be a more reliable cue to SoA. Therefore, the effects of selection fluency on SoA should decrease across trials (pattern **b**).
3. *Generally prospective.* Selection fluency is a heuristic that is generally used to prospectively inform SoA, independent of outcome monitoring. In this case, the effects of selection fluency on SoA would remain constant across trials (option **c**).
4. *Prospective effects on learning.* Selection fluency impacts on the learning of action-outcome associations. As fluent actions are well linked to their outcomes, agency ratings increase steeply, whereas dysfluent actions are

more slowly associated with their outcomes. This account predicts that the effects of selection fluency on SoA would increase across trials (pattern **a**).

Notably, accounts 1 and 4 predict the same pattern of results (**a**), but for very different reasons. Whilst these may be prove difficult to dissociate, it is worth noting that account 4 would be compatible with using selection fluency as a heuristic cue to agency in different contexts. In contrast, account 1 would not be compatible with this heuristic, since it implies that selection fluency is exclusively learned in a context-specific manner.



**Figure 5.1.** Rationale for the study. **A.** Prospective and retrospective cues to the SoA, with their presumed corresponding variables investigated here (in *italics*). **B.** Schematic of hypothetical results for the interaction between selection fluency and trial number, assuming that repeated exposure to actions and outcomes will influence SoA. As knowledge about action-outcome contingencies is gathered (across trials), the effect of selection fluency on SoA might: a) increase; b) decrease; or c) remain constant.

Finally, it is worth considering the role of AOI. This was manipulated in the experiments considered, in order to increase variability in SoA, and prevent ceiling effects on ratings (Haggard et al., 2002; Wenke et al., 2010). No significant

interactions have been found between selection fluency and AOI so far. However, the relation between AOI and outcome identity is less clear. Longer delays are typically associated with a weaker SoA (Farrer et al., 2013), possibly due to short AOIs being strongly associated with the typical windows of motor control. Therefore, short intervals may generally lead to stronger SoA than long intervals, regardless of knowledge about outcome identity. Yet, AOI may also compete with outcome identity in driving SoA. A focus on outcome identity, as instructed in Group 1, may reduce the contribution of AOI to SoA. In contrast, if contextual cues highlight using AOI as a cue, as in Group 2, outcome identity may become less relevant to SoA. This would predict a reduced influence of outcome identity on SoA, and thus a smaller change in agency ratings across trials.

## **5.2. Methods**

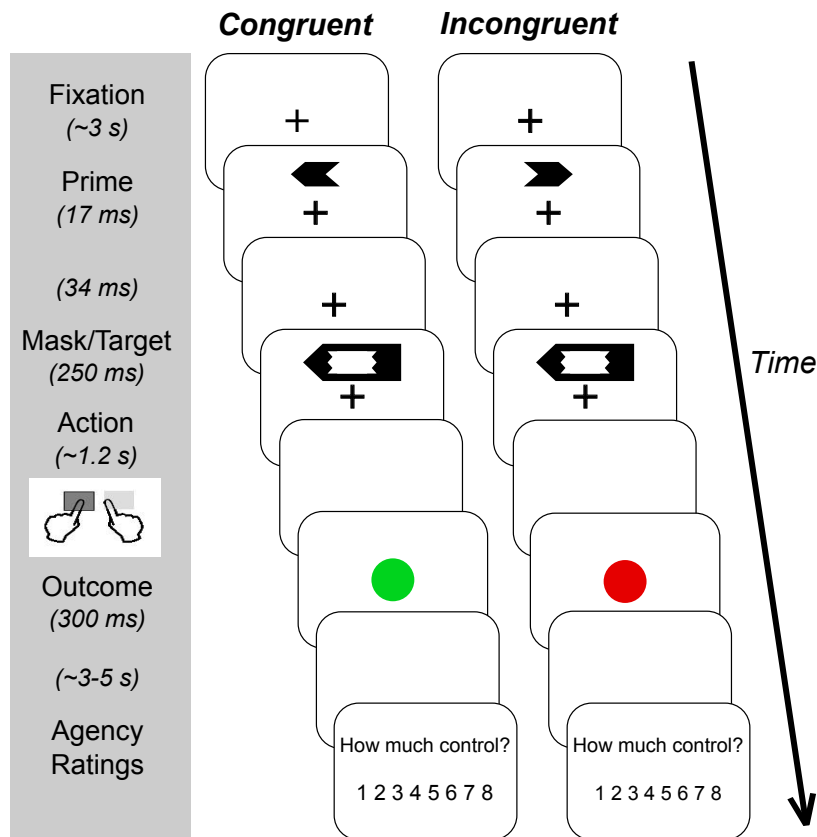
### **5.2.1. Experimental Design**

The behavioural tasks are described in detail in each respective publication; the common design features are depicted in **Figure 5.2**. A list of the studies included is provided in **Table 5.1**. This table further lists the grouping of experiments, and relevant differences between the studies. In addition to the differences listed, experiments in Group 1 used a 9-point Likert scale, whereas studies in Group 2 used an 8-point scale.

### **5.2.2. Study Selection for Multi-Study Analysis**

Of all the known studies that have manipulated action selection and obtained a measure of agency, three were excluded from our analysis. One study only obtained agency judgements at the end of each block (Wenke et al., 2010). Another study only had two outcomes, therefore did not require much action-outcome learning (Damen et al., 2014). A final study could not provide detailed trial-wise data due to technical error during data collection (Chambon & Haggard, 2012). Furthermore, we only included the relevant data from the available studies: For Exp. 2, we only used the data for the half of participants who gave agency ratings at the end of each trial,

as the other half gave block-wise ratings. Also, we excluded trials from the neutral condition. For Exp. 6, we only used data from healthy participants, as the others were patients. For Exp. 7, we only used the data from the sham TMS (i.e. control) condition, since the other conditions involved active TMS stimulation.



**Figure 5.2.** Task outline for the subliminal priming studies (adapted from Chambon et al., 2013). This outline is similar to experiments from Group 1 (see Chapters 2 & 4). Critically, all studies involve a target which calls for a left- or right-hand action. This is followed by a coloured circle (the outcome), after a variable delay. Participants give agency ratings over the outcome, at the end of each trial. In studies with supraliminal flanker, targets consist of one of two letters, which appear surrounded by congruent (e.g. SSSSS), or incongruent (e.g. HHSHH) flankers.

### 5.2.3. Data Analysis

Our main goal was to model the trial-wise data obtained across the seven experiments in a three-level multilevel linear regression model (e.g. Gelman & Hill, 2006), and explore the relative contributions of selection fluency, AOI and outcome identity to SoA. Given the difference in instructions, we further considered the effect of Group. More details are provided below.

**Table 5.1.** Factorial design and differences between studies.

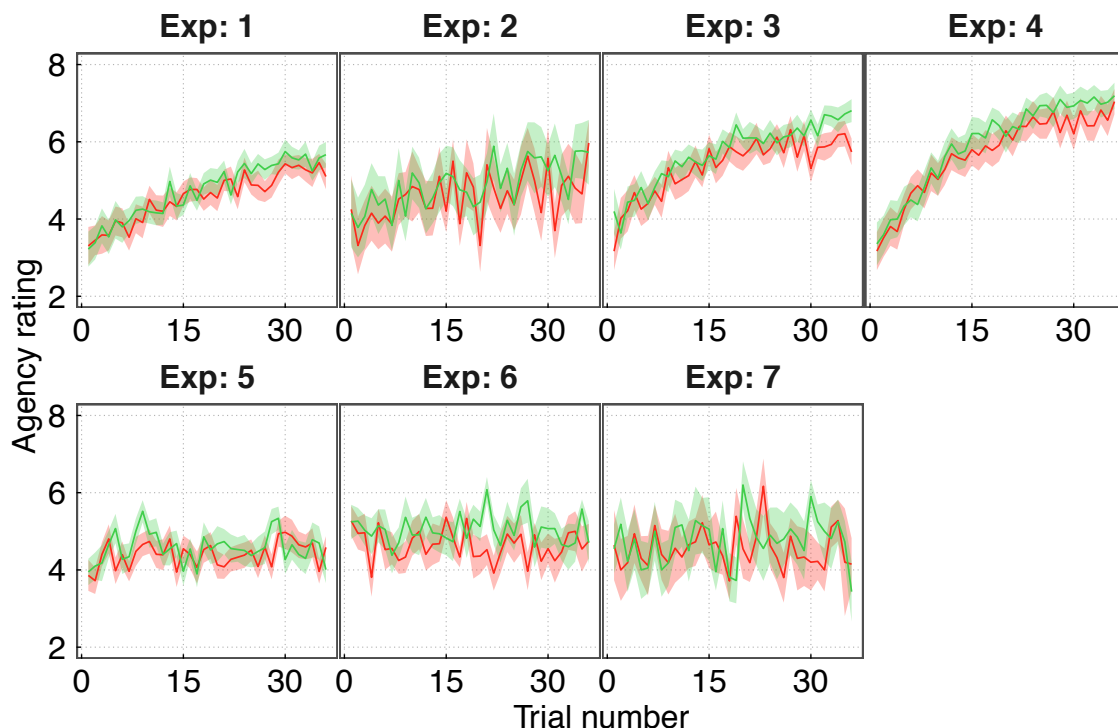
Publication	Group	Experiment Number	Manipulation	Action-Outcome Interval (ms)	Number of Outcomes <sup>1</sup>	Number of Trials	Number of Blocks
Chapter 4	1	1	Subliminal Priming	400, 600	8	64	8
Chapter 2, Experiment 1	1	2	Supraliminal Flankers	100, 300, 500	4	72	4
Chapter 2, Experiment 2	1	3	Supraliminal Flankers	200, 400	8	64	8
Chapter 2, Experiment 3	1	4	Supraliminal Flankers	100, 500	4	64	12
Chambon et al., 2013	2	5	Subliminal Priming	100, 300, 500	4	48	4
Chambon et al. (Unpublished)	2	6	Subliminal Priming	100, 400, 700	8	56	8
Chambon, Moore, et al., 2014	2	7	Subliminal Priming	100, 400, 700	4	36	2

<sup>1</sup> This refers to the number of outcomes (coloured circles) per block. In Experiment 4, there were 12 coloured circles overall, but only 4 in each block.

To detect potential differences between experiments, and groups, we plotted the data for each experiment separately (see below). We also estimated the effects in separate multilevel models for each experiment, including all trials from each experiment, confirming the results presented here (see **Appendix D**).

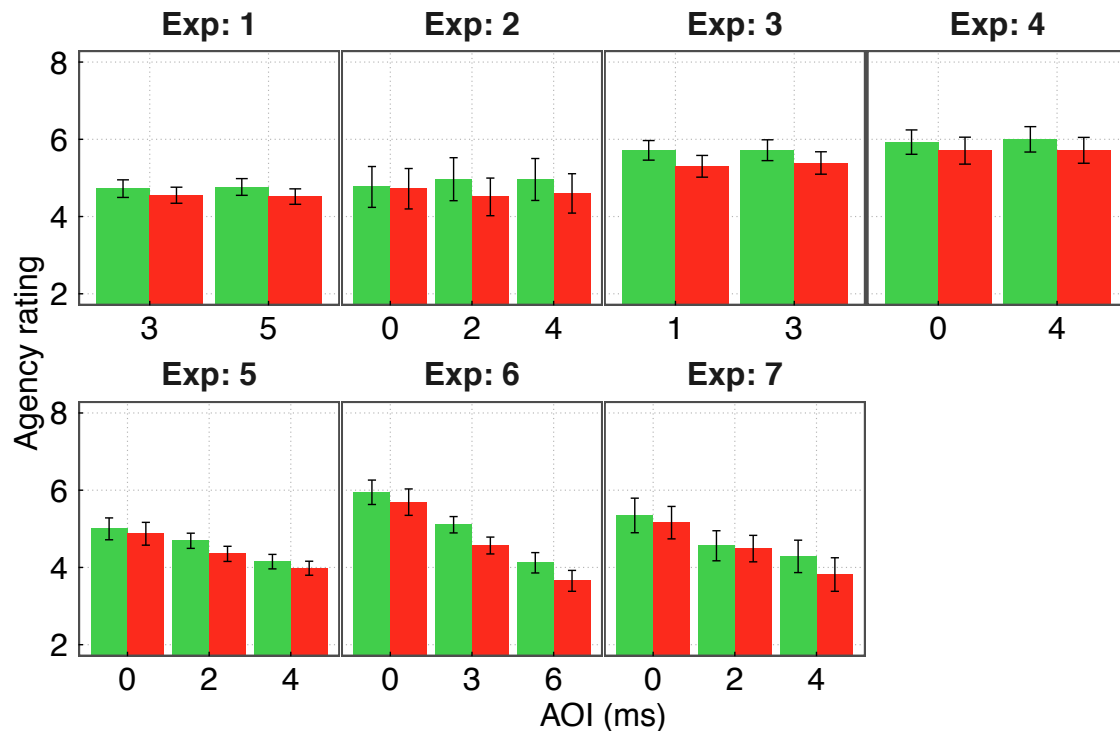
### 5.3. Results

Considering how agency ratings changed across the trials, the data presented in **Figure 5.3** suggest an important distinction between the patterns of data obtained in the two groups of experiments. In Group 1 (Exp.: 1-4), agency ratings increased over trials, and the effect of congruency (higher agency ratings for congruent trials) increased as participants progressed through the trials, within a block. These patterns seemed largely absent in the data obtained from experiments in Group 2 (Exp.: 5-7): Agency ratings and the congruency effect remained broadly stable throughout the block.



**Figure 5.3.** Mean agency rating across participants as a function of trial number, for each experiment ( $\pm 1$  SEM as shaded area). Top row = experiments from Group 1, bottom row = Group 2. Green = congruent, red = incongruent.

We also considered the effect of AOI on agency ratings. The data presented in **Figure 5.4** suggest another important difference in the patterns of data between the groups of experiments: Longer AOIs led to lower agency ratings in Group 2, but this pattern seemed to be absent in Group 1.



**Figure 5.4.** Mean agency rating as function of AOI for each experiment ( $\pm 1$  SEM). Top row = experiments from Group 1, bottom row = Group 2). Green = congruent, red = incongruent.

To formally quantify these patterns in the data, we used a three-level multilevel regression model, with trials nested as repeated measures within individuals, and individuals nested within experiments. Before modelling the data, we limited the block length of each experiment to 36 trials per block in order to prevent any effects of trial number due to differences in block length across the experiments included in the multi-study model (see **Table 5.1**).

### 5.3.1. Multilevel Model of Agency Ratings

We modelled these data with a three-level multilevel linear regression model (e.g. Gelman & Hill, 2006), where the effects of trial (linear and quadratic), congruency,



AOI, congruency by trial (linear) interaction, and congruency by AOI varied within individuals and experiments. We also modelled these effects at the average level, in addition to their interactions with a group indicator variable. The full average effects specification of this model, along with estimated effects is shown in **Table 5.2**. Our main goals of inference were the differences in the data patterns observed in **Figure 5.3** and **Figure 5.4**; these effects are highlighted in bold in **Table 5.2**.

The use of multilevel modelling was motivated by acknowledging the hierarchical structure of the data, with repeated measures on individuals, and individuals within experiments, and the need to include continuous predictors at all levels of analysis. Additionally, the data were not balanced across individuals or experiments, because the sequence of congruent and incongruent trials within a block was randomized, further motivating the use of multilevel modelling. The model was fitted using the *lme4* package in the R statistical programming environment (Bates et al., 2014; R Core Team, 2015).

We used the following coding scheme for the predictor variables in the regression model: Group was dummy coded as 0 (Group 1) and 1 (Group 2), therefore all effects in **Table 5.2** without Group interaction denote average effects for experiments in Group 1. Trial was entered as a linear and quadratic predictor (computed using orthogonal polynomials, range of linear trial predictor = [-0.28, 0.28]); Congruency was coded as -0.5 (incongruent) and 0.5 (congruent); AOI denotes the effect of 100ms increase in action-outcome interval, and was centred at 400ms and used as a linear predictor—when using AOI as a factorial predictor, the model did not converge, and the data presented in **Figure 5.4** suggest that a linear effect in these ranges of AOI was a good approximation.

Regarding the main effect of congruency, the model (see **Table 5.2**) showed a strong and robust effect of congruency on SoA for Group 1 (*main effect of Congruency*), with no noticeable difference in this effect for Group 2 (*Congruency by*

*Group interaction*). AOI did not exert a detectable effect in Group 1 (*main effect of AOI*), but strongly decreased agency ratings in Group 2 (*AOI by Group*, general linear hypothesis test<sup>2</sup> of AOI effect for Group 2:  $b = -0.27$ ,  $t_{(7)} = -7.91$ ,  $p < .001$ ; see **Figure 5.4**). Importantly, there was no significant interaction between AOI and congruency for Group 1 (*Congruency by AOI*), nor for Group 2 (*Congruency by AOI by Group*).

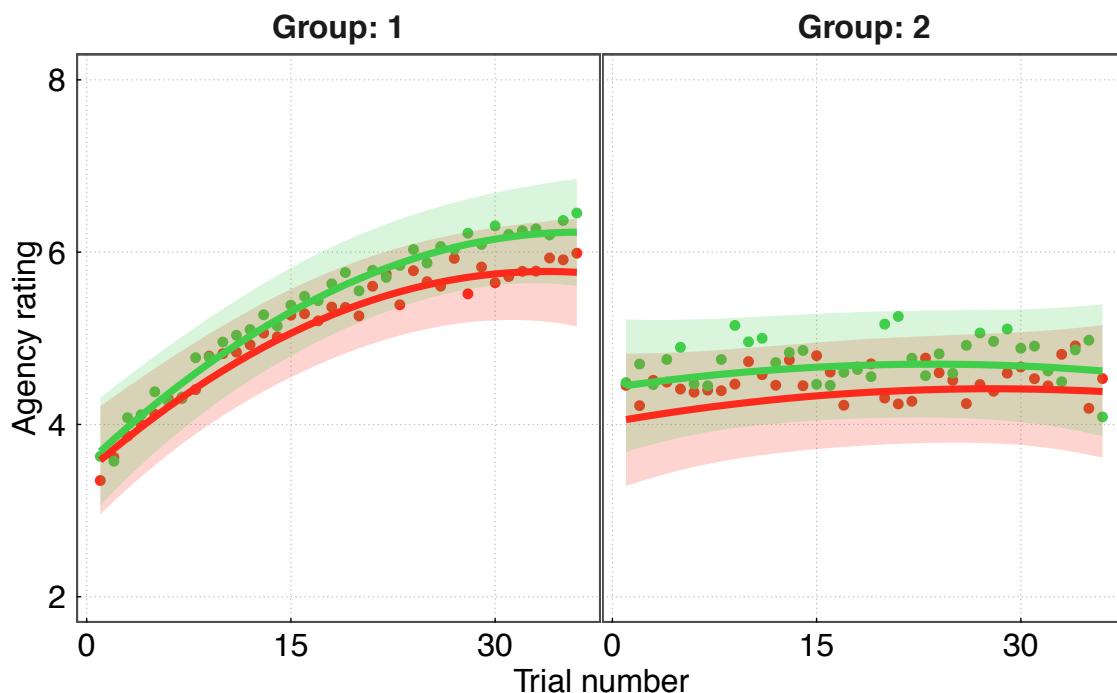
**Table 5.2.** Multilevel model of agency ratings. Effects of primary interest are highlighted in bold.

Average Effect	Estimate (SE)	$t$	$p$	95% CI	
				Lower	Upper
Intercept	5.22 (0.25)	20.80	< 0.001	4.73	5.72
<b>Trial</b>	<b>4.21 (0.57)</b>	<b>7.32</b>	<b>0.001</b>	<b>3.08</b>	<b>5.34</b>
Trial (quadratic)	-1.18 (0.31)	-3.82	0.014	-1.79	-0.58
<b>Congruency</b>	<b>0.28 (0.05)</b>	<b>5.25</b>	<b>0.008</b>	<b>0.18</b>	<b>0.39</b>
<b>AOI</b>	<b>0.01 (0.03)</b>	<b>0.43</b>	<b>0.665</b>	<b>-0.04</b>	<b>0.06</b>
Group	-0.75 (0.39)	-1.90	0.110	-1.52	0.02
<b>Trial x Congruency</b>	<b>0.65 (0.19)</b>	<b>3.42</b>	<b>0.001</b>	<b>0.28</b>	<b>1.02</b>
Congruency x Group	0.03 (0.10)	0.34	0.746	-0.16	0.23
<b>AOI x Group</b>	<b>-0.28 (0.04)</b>	<b>-6.60</b>	<b>&lt; 0.001</b>	<b>-0.37</b>	<b>-0.20</b>
Congruency x AOI	0.02 (0.02)	1.05	0.293	-0.01	0.05
<b>Trial x Group</b>	<b>-3.76 (0.91)</b>	<b>-4.14</b>	<b>0.008</b>	<b>-5.55</b>	<b>-1.98</b>
Trial (quad.) x Group	0.90 (0.50)	1.79	0.125	-0.08	1.88
<b>Trial x Congruency x Group</b>	<b>-0.92 (0.39)</b>	<b>-2.36</b>	<b>0.019</b>	<b>-1.69</b>	<b>-0.16</b>
Congruency x AOI x Group	0.02 (0.03)	0.67	0.500	-0.04	0.08

Note.  $p$ -values were computed using the Satterthwaite approximation (Kuznetsova et al., 2015). Confidence intervals were obtained using the Wald method (Bates et al., 2014). 'x' denotes interaction terms.

<sup>2</sup> Obtained using the model's estimated parameters and variance-covariance matrix. All tests were computed with a conservative 7 degrees of freedom (number of experiments).

Regarding the effect of outcome identity, agency ratings increased with trial number in Group 1 (*main effect of Trial*), but not in Group 2 (*Trial by Group*, general linear hypothesis test of Trial effect for Group 2:  $b = 0.45$ ,  $t_{(7)} = 0.63$ ,  $p = .55$ ). The effect of congruency increased with trial number in Group 1 (*Trial by Congruency*), but this interaction was not noticeable in Group 2 (*Trial by Congruency by Group*, general linear hypothesis test of Trial by Congruency for Group 2:  $b = -0.28$ ,  $t_{(7)} = -0.81$ ,  $p = .45$ ). These effects are illustrated in **Figure 5.5**.



**Figure 5.5.** Multilevel model of agency ratings. Predicted average agency ratings for each Group as a function of trial number (with 95% prediction intervals as shades around regression lines; dots are means from raw data across experiments within each Group). Green = congruent, red = incongruent. Predictions were obtained from 100000 simulations from the posterior distribution of plausible parameter values under uniform priors (Gelman & Su, 2015). Note the progressively increasing effect of congruency on agency in studies of Group 1, but not of Group 2.

## 5.4. Discussion

The present study investigated the relative contribution of prospective and retrospective cues to agency. Overall, a multi-study analysis of seven experiments showed a robust effect of the prospective – selection fluency – cue to SoA. These

effects were consistent across different instructions regarding the agency ratings procedure, suggesting a general role of this prospective cue to agency. The different instructions did, however, modulate the contribution of the two retrospective cues – outcome identity and AOI. On the one hand, Group 1 experiments gave instructions to focus on learning action-outcome contingencies. These showed an increase in agency ratings across trials, as knowledge about the outcomes was gathered. Yet, no effect of AOI was found, suggesting that outcome identity drove agency ratings for Group 1. On the other hand, Group 2 experiments instructed participants to additionally attend to AOI. These did not show a consistent change in ratings across trials, but did show effects of AOI on ratings. Thus, for Group 2, AOI drove agency ratings, instead of outcome identity. These results suggest a trade-off or competition between the two retrospective cues. As it is our main interest, we will first consider the relation between prospective and retrospective cues.

### **5.4.1. Prospective vs. Retrospective Cues**

For Group 1 studies, which focused on outcome identity, interactions with the prospective cue emerged. In addition to agency ratings increasing over time, we found a concurrent increase in selection fluency effects. There was a shallower increase over time in agency ratings for dysfluent actions (incongruent priming/flankers), relative to fluent actions (congruent priming/flankers). For Group 2 studies, which focused on AOI, there were no interactions between retrospective and prospective cues to agency. As there was no change in agency ratings across trials, selection fluency effects remained stable throughout. Dysfluent actions (incongruent priming) led to lower agency ratings than fluent actions (congruent priming) across trials. We will now evaluate the 4 possible mechanisms outlined in the introduction with regards to our results.

#### **5.4.1.1. Learning to be Prospective: Fluency Effects Are Not Context-Specific**

The interaction between selection fluency and outcome identity observed for Group 1 studies are in line with the prediction of the “learning to be prospective” account. This view proposes that we learn to use selection fluency as a cue to agency in a context-

specific manner: We learn that fluent actions are associated with specific outcomes, and come to rely on selection fluency as a prospective proxy for the action outcome. From this perspective, there should be no fluency effects at the start of a block, because there is no *a priori* knowledge about how fluency informs SoA (hypothesis **a**), in **Figure 5.1**). This was indeed the pattern observed in Group 1 studies (see **Figure 5.5**).

However, this account cannot explain the general effects of selection fluency found in Group 2 (see **Figure 5.5**). In those studies, similar fluency effects were present from the start of a block, and independently of outcome identity. Therefore, the relation between selection fluency and SoA had to be learnt in advance. Note that this does not preclude the possibility that the association between selection fluency and action outcomes could still be strengthened in particular contexts, and further enhance the contribution of prospective cues to SoA. For example, in complex tasks, e.g. sports, as expertise increases, greater fluency in action selection will typically be associated with greater accuracy in outcome prediction (Gray et al., 2007). Thus, the experience of fluency might gradually become an even more reliable advance predictor of action outcomes, and agency.

#### 5.4.1.2. Optimally Prospective: No Reduction in Fluency Effects

The present results seem to rule out the “optimally prospective” account. Assuming prior knowledge about the relation between selection fluency and SoA, this account suggests that we rely most on selection fluency as a cue to agency when other cues are unreliable/unavailable. An optimal cue integration account (Moore & Fletcher, 2012; Synofzik et al., 2013) would have predicted that the prior reliability of the selection fluency cue could have served to compensate for the low reliability of outcome identity at the start of a block, and perhaps that its contribution to SoA might reduce with outcome learning, given that agency ratings should have been tracking the outcomes (hypothesis **b**), in **Figure 5.1**). In contrast, results showed either a constant or an increasing contribution of selection fluency to SoA across trials. These findings question the use of an optimal integration across *all* cues to SoA (Sidarus et al., 2013), although it may apply when integrating *some* cues. Current proposals do

not easily account for changes in reliability over time, such as those found for outcome identity during instrumental learning. Moreover, it remains unclear how these processes can handle the integration of cues that become available at different times *within* a trial, e.g. prospective vs. retrospective.

#### **5.4.1.3. Generally Prospective: Selection Fluency as a General Heuristic**

These findings clearly support the “general heuristic” account described above: Selection fluency can serve as a heuristic cue to agency, which can be employed in novel circumstances, independently of one’s knowledge about action-outcome contingencies (hypothesis **c**), in **Figure 5.1**). In Group 2 studies, we did not find a general increase in agency ratings, as predicted by hypothesis **c**), but, importantly, selection fluency effects were similar across trials.

These findings suggest that the relevance of selection fluency to agency is likely learned through everyday experience, rather than being specifically linked to any given environment. Regardless of a specific link between fluency/difficulty and particular outcomes, fluent action selection is more likely associated with desired or predicted outcomes than dysfluent selection. These findings are also consistent with the view that response conflict is an aversive signal (Botvinick, 2007), with negative affective consequences that can affect subsequent events (Fritz & Dreisbach, 2013). Relatedly, many studies have shown general influences of fluency/difficulty, e.g. in stimulus processing, on a variety of judgements, such as liking or familiarity (Alter & Oppenheimer, 2009).

#### **5.4.1.4. Prospective Effects on Action-Outcome Learning**

The increased effects of selection fluency on agency across trials seen in Group 1 would also be consistent with the “prospective effects on learning” account. From this perspective, outcomes that follow dysfluent actions would be less easily associated with their corresponding action, relative to outcomes that follow fluent actions. Assuming agency ratings would be tracking action-outcome knowledge, ratings would increase more slowly for outcomes that followed dysfluent actions, thus the difference in ratings relative to outcomes that followed fluent actions would

increase. This could potentially result from action representations being disrupted by response conflict, and, in turn, be less effectively associated with outcomes. It may be adaptive to learn less about the consequences of dysfluent actions, than of fluent actions. For example, a novice playing darts who hits the bull's-eye at the first throw should recognise that this successful hit may have been partly due to luck. Her lack of practice with the game means her action selection was likely not very fluent or precise, and she is unlikely to easily replicate such an ideal hit.

In fact, sensorimotor predictions can be influenced by selection fluency (Stenner, Bauer, Heinze, et al., 2014), and thus alter action-outcome linkage. Otherwise, the negative affect induced by conflict could also impair the associative process. Interestingly, this view would suggest that selection fluency effects previously found with block-wise agency ratings (e.g. Wenke et al., 2010) could result from better or worse associative links between particular outcomes and their corresponding actions

Finally, under the instructions to focus on outcome identity, agency ratings may have been very low at the start of a block, since participants did not have any outcome information yet. It is possible that this floor effect could have masked the general, and typically stable, influence of selection fluency on agency, which then gradually emerged as outcome knowledge was gathered, and ratings started to vary. Further research is needed to test these hypotheses, namely by directly testing knowledge about action-outcome associations.

### 5.4.2. Competition Between Retrospective Cues

Returning to the relation between the two retrospective cues, outcome identity and AOI, the apparent competition, or trade-off, between the two cues could perhaps be partially attributed to differences in the salience of the AOI cue. Variability may have been less salient in most experiments that were part of Group 1, as three out of four used only 2 intervals, whereas experiments in Group 2 always used 3 intervals (see **Table 5.1**). However, similarly to those experiments, and others (Chambon & Haggard, 2012; Sidarus et al., 2013), 3 AOIs were used in Experiment 2 (Group 1), but no significant effect of AOI was found (see **Table D2**, in **Appendix D**). Therefore,

the difference in instructions between the groups of studies likely played a greater role in the relative weighting of the retrospective cues.

When participants were instructed to use AOI as a cue, and given the prior association between short intervals and stronger SoA, this may have seemed a more reliable cue to SoA, relative to outcome identity. The availability and relevance (reinforced by the instructions) of this cue to agency could have, in turn, reduced the learning of action-outcome associations. It is worth noting that the learning of action-outcome associations in these studies could be difficult, even when only 4 colours appeared in each block (as in Exps. 5 & 7). In the subliminal priming paradigm, participants could learn that 2 colours were associated with each hand, but could not further disambiguate those 2 colours. Outcome colours depended both on the action performed, and on prime-action congruency, and this latter information was not available since primes were subliminal. On the other hand, in studies with supraliminal flankers, congruency information was explicitly available, which could have facilitated action-outcome learning. Whereas all studies in Group 2 used subliminal priming, and 3 of 4 studies in Group 1 used supraliminal flankers (see **Table 5.1**), one might wonder whether differences in learning effects across groups of studies could explain the different role of outcome identity to SoA. However, in Experiment 1 (Group 1) an increase in agency ratings over time was found (see **Table D1**) despite using subliminal priming, and having eight outcome colours per block (due to another factor in the experiment). This further supports a role for instructions.

Nonetheless, while this trade-off may emerge due to instructions, other contextual cues could also have an impact. A study which used similar instructions to those in Group 2 reported both an effect of AOI and of outcome identity on agency ratings (Sidarus et al., 2013), with no significant interaction. Moreover, an interaction between outcome identity and selection fluency was found. However, this study involved a lower action-outcome contingency (67%), than the studies considered here (100%). As outcome predictions were sometimes violated, a stronger intrinsic motivation to attend to outcome identity may have been engaged. This motivation



may have been partially absent in the studies of Group 2 investigated here, as outcome identity was fully contingent on action. Alternatively, if outcome identity is highly reliable, given prior training, this may be a predominant cue to SoA (Evans, Gale, Schurger, & Blanke, 2015). Therefore, interactions among retrospective, or between retrospective and prospective cues to agency, can be contextually modulated not only by instructions, but also by the requirements of the task at hand.

### **5.4.3. Conclusions**

The present investigation has shown that action selection fluency can serve as a heuristic cue to prospectively inform our SoA. The experience of agency may already begin even before the action. This prospective SoA may serve as a general advance predictor of successful action, and to bridge the interval between action and outcome. Importantly, the SoA requires a complex integration of multiple cues, from multiple sources, available at different times. Prospective and retrospective cues can have an independent effect on SoA. Nonetheless, depending on contextual cues, such as instructions, or task requirements, the relative contribution of these two cues to SoA may also be dynamically changed during instrumental learning. We speculate that dysfluent action selection may weaken the link between action and subsequent outcome.



## **Chapter 6. Beyond Self-serving Bias: Diffusion of Responsibility Reduces Sense of Agency and Outcome Monitoring.**

*Diffusion of responsibility has been proposed to underlie decreased helping and increased aggression in group behaviour. However, direct effects of the presence of other people on how we experience the consequences of our actions have not been shown so far. In this EEG-study, we investigated whether diffusion of responsibility reflects a purely post-hoc self-serving bias, or has online effects on how we process outcomes of our own actions, and how we experience sense of agency with respect to them. Using a task in which objective responsibility for an outcome was unambiguous, we show that the alleged presence of another agent reduced participants' subjective sense of agency over the outcomes of their own actions. Furthermore, amplitude of the feedback-related negativity evoked by outcome stimuli was decreased, suggesting reduced monitoring of action outcomes. The presence of other agents may lead to diffused responsibility by modulating how the brain relates actions to outcomes.*

## 6.1. Introduction

Social psychology has long recognised that the presence of other people can have substantial effects on individual behaviour. One example is the so-called “bystander effect”: in emergency situations requiring the help of a bystander, the likelihood of someone offering that help decreases with the number of people witnessing the emergency (Darley & Latane, 1968). People are also less likely to react to social norm violations, such as spraying graffiti, when bystanders are present (Chekroun & Brauer, 2002). A similar effect, termed “social loafing”, can be observed in working environments. When a group of people has to work towards a collective goal, each individual on average puts in less effort than they would when working alone (Karau & Williams, 1993). Experimental studies also show that groups tend to make riskier choices than individuals (Bradley, 1995; Wallach, Kogan, & Bem, 1964) and behave more aggressively (Bandura, Underwood, & Fromson, 1975; Meier & Hinsz, 2004). All these situations have in common that individual behaviour is altered by the presence of other people. In such experiments, participants report that they feel less responsible for the outcome of group decisions, especially those with negative consequences (Forsyth, Zyzniewski, & Giammanco, 2002; Mynatt & Sherman, 1975).

These findings have led to the concept of diffusion of responsibility: the idea that the presence of others changes the behaviour of the individualy making them feel less responsible for the consequences of their actions (Bandura, 1991).Diffusion of responsibility has even been proposed to serve moral disengagement when committing inhumane actions (Bandura, 1999). However, it is not clear to date whether diffusion of responsibility actually has a causal effect on behaviour and experienced responsibility, or merely constitutes a post-hoc bias in self-reports, serving the preservation of self-esteem and impression management. In order to play a causal role in group behaviour, diffusion of responsibility would need to have online influences on how people experience a given situation. The mechanisms by which this influence could occur remain unclear.

A central aspect of human experience is the sense of agency, i.e. the feeling that we can control external events through our actions. This experience depends on establishing a relation between our actions and their consequences. When the outcomes of our actions do not match our predictions or expectations, we feel a loss of agency (Blakemore et al., 2002). The sense of agency plays a crucial role in social interactions, where authorship of events can be ambiguous (Frith, 2014). Moreover, sense of agency is tightly linked to the experience and allocation of responsibility: we should only feel responsible for events that we had at least some control over.

The subjective experience of agency can be measured by asking people how much control they felt over a certain outcome of their actions. However, similarly to diffusion of responsibility studies, such self-reports can be subject to biases and post-hoc considerations. A more objective, and implicit, measure of action-outcome processing can be obtained using event-related potentials (ERPs). The feedback-related negativity (FRN) is an ERP-component observed when participants receive negative feedback (for a review, see San Martín, 2012). Importantly, this component is sensitive to the perceived controllability of action outcomes: when participants believe that an outcome is uncontrollable, the FRN to negative outcomes is greatly reduced (Li et al., 2011; Yeung et al., 2005). As the FRN is sensitive to the motivational significance of outcomes (Gehring & Willoughby, 2002; Holroyd & Yeung, 2012), a likely explanation for this finding is that uncontrollable outcomes are less important to the agent, as they provide little information on how to improve future behaviour.

One way in which the presence of others may reduce sense of agency is through increased authorship ambiguity and an objective decrease in control: a communal grade for a group project provides little information about the quality of individual contributions. Accordingly, Li et al. (2010) showed that in a dice-tossing task, FRN amplitude was reduced when instead of tossing all three dice, participants tossed only one, and the others were tossed by other players. Playing with others also reduced participants' self-reported contribution to the outcome.

However, diffusion of responsibility occurs beyond an objective distribution of control among agents. For example in the helping scenarios described for the bystander effect, each individual's control over their action outcomes is not reduced, nor is ambiguity of attribution increased. If several people witness an emergency and one person decides to act, their actions have the same consequences as if they were alone. Thus, to explain why the presence of others changes people's behaviour, diffusion of responsibility would have to influence an individual's experience of the situation, beyond objective effects on action-outcome contingencies. To date, it remains unclear whether people actually experience the link between their own actions and outcomes as being weakened by the presence of other potential agents.

We propose that if diffusion of responsibility causally affects group behaviour, it does so by reducing people's sense of agency over action outcomes in the presence of others. This reduction in sense of agency may be mediated by the complexity of social decision making compared to individual decision making. In social situations, we need to take into account the possible actions of others, and the outcomes those actions could produce, in addition to our own actions and outcomes. Difficulty, or dysfluency, in decision making has been shown to lead to a reduction in sense of agency over action outcomes (for a review, see Chambon, Sidarus, et al., 2014). Therefore, the increased difficulty of decision making in social contexts could potentially contribute to reduced sense of agency. Importantly, it should do so in a situation where attribution of outcomes to one's own actions is unambiguous, and action-outcome contingencies are identical for social and non-social settings.

To this end, we designed the task to create two experimental conditions that only differed in terms of social context. This required the following criteria to be met: 1. Action consequences had to be controllable, i.e. there had to be a consistent contingency between actions and outcomes; 2. While controllable, outcomes should not be perfectly predictable, to ensure that outcome feedback was relevant to the participant on each trial; 3. Attribution of outcomes to the participant's own actions had to be unambiguous in both the social and non-social context.

To meet these requirements, we used a task in which a marble rolls down a bar, and an action is required to stop it from crashing (Schel, Brass, Haggard, Ridderinkhof, & Crone, 2014). In our version of this task, participants either played alone, or (in their belief) together with another player. Importantly, the stopped marble provided participants with immediate feedback of their action, thus eliminating ambiguity as to who caused a given outcome. Stopping the marble incurred some cost for the participant, but this cost was avoided if the other player acted to stop the marble. However, if the marble crashed the cost was much greater. ERPs were recorded in response to the feedback of points lost, and participants rated how much control they felt over that outcome.

As diffusion of responsibility is mostly used to explain behaviour in situations where acting is somehow costly or effortful, or results in negative consequences, we designed the task to exclusively produce negative outcomes. If diffusion of responsibility is merely a post-hoc justification, this loss frame would provide ideal conditions for a self-serving bias to “attribute away” responsibility to other agents. In that case, participants should report especially low sense of control in social trials with worse outcomes.

We predicted that the presence of an alternative agent would result in a diffusion of responsibility effect. This might result in changes in behaviour, as participants might wait for the other player to stop the marble, and a reduction in the sense of agency. Furthermore, if diffusion of responsibility reduces the subjective experience of control over action outcomes, the FRN component to the outcome should be reduced.

## **6.2. Materials and Methods**

### **6.2.1. Participants**

We tested 32 healthy volunteers (16 male, 16 female; age 18-32). The data of one participant was lost due to technical failure. The data of 6 other participants was excluded from data analysis based on the following a priori criteria: fewer than 20 trials per condition after artefact rejection of ERP epochs (3 participants);

spontaneously expressing suspicion about the co-player participation in the task in the post-experimental questionnaire (3 participants). Thus, data of 25 participants (12 male, 13 female) was included in the analyses.

### **6.2.2. Apparatus and Materials**

Participants were tested in pairs. Stimuli were presented on two identical computer screens for the two participants. Participants gave responses using standard computer mice. After the task, participants filled out a post-experimental questionnaire probing for suspicion concerning the participation of the co-player in the task, as well as the Locus of Control Scale by Rotter (1966), and the subscales “Diffuse Responsibility” and “Exercised Responsibility” of the Ascription of Responsibility Questionnaire (Hakstian, Suedfeld, Ballard, & Rank, 1986).

EEG was recorded from 26 channels using g.tec g.USB amplifiers with active ring electrodes and non-abrasive conductive gel. Horizontal and vertical eye movements were recorded simultaneously. EEG signals were referenced online against the left earlobe and were recorded with a 0.1 Hz highpass filter.

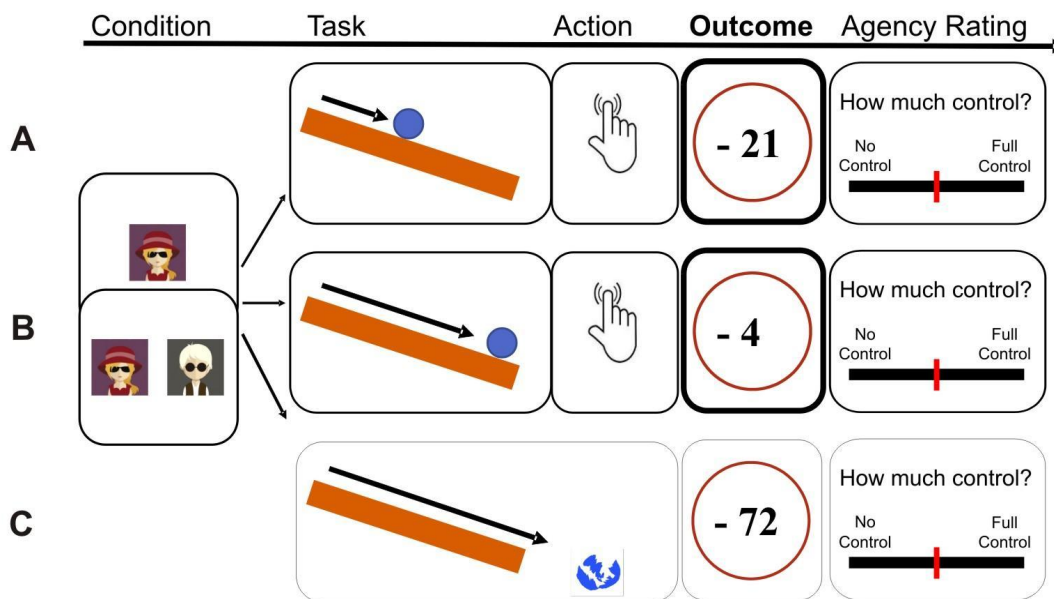
### **6.2.3. Design and Procedure**

Participants were invited to the laboratory in mixed-gender pairs of two. They received instructions together, filled out consent forms for participation in the study and were then seated in adjoining laboratories for the testing. During the instructions, participants were assigned one avatar, which would represent them during the task. They were also shown their co-player’s avatar, which would be used when they played together. Both participants performed the task simultaneously, but separately. After the task was finished, participants filled out post-experimental questionnaires and personality questionnaires (see materials above). Participants were then fully debriefed and paid for their participation. Payment consisted of £7.50 per hour, plus any earnings from the task.

The marble task was designed to create a situation in which acting was costly, but withholding action was potentially more costly still. In each trial, participants had to



stop a rolling marble from falling off a tilted bar, and crashing (see **Figure 6.1**). Participants were instructed that, at the beginning of each block, they would receive 1500 points worth 150 pence, and in each trial they could lose up to 100 of these points. The task consisted of 4 blocks of 30 trials each. Trials were randomly assigned to either the “Alone” or the “Together” condition, with 15 trials per condition and block.



**Figure 6.1.** Marble Task. This figure shows the outline of a low-risk successful trial (A), a high-risk successful trial (B), and an unsuccessful trial (C). Note that C is the worst outcome, B the best, and A the intermediate. Social context was indicated at the start of a trial, by either presenting the participant’s own avatar alone, or together with the other player’s avatar. The marble colour served as a reminder of social context, and was either blue in the alone condition (shown here), or green in the together condition. In the together condition, besides the trials displayed here, there were trials in which the “other” player stopped the marble, and the participant did not lose any points. ERPs were time-locked to outcome presentations of successful trials (A and B, marked in bold) in which the participant stopped the marble.

In the beginning of an “Alone”-trial, participants saw their own avatar alone, indicating they would be playing by themselves, while their co-player supposedly played simultaneously on his/her computer. Next, they saw a blue marble lying on top of a tilted bar, which after 500ms started rolling down towards the lower end of the bar. At any point, participants could press the left mouse button to stop the marble. If they did so, the marble stopped in its current position, providing immediate

feedback of their successful action. If participants did not react in time, the marble rolled off the bar and crashed. The final position of the marble, whether stopped or crashed, was shown for 500ms, followed by the presentation of a fixation cross for 1500 – 2500ms. In either case, participants received information about how many points they lost, i.e. the action outcome for 2000ms. ERPs were time-locked to outcome presentation. Afterwards, a fixation cross was presented for 500ms and then participants saw a visual analogue scale with the question “How much control did you feel over the outcome?” and the end points of the scale labelled “No control” and “Complete control”. Participants used the mouse to indicate how much control they felt they had over the number of points lost during that trial.

Participants were instructed that the later they stopped the marble, the fewer points they would lose. In order to make it difficult to always stop the marble at the very end of the bar, the speed with which the marble rolled down the bar varied from trial to trial. Also, at some point along the bar, the marble would speed up, and this point varied from trial to trial. This added a risk component to the task, since if the participant waited too long, the marble might suddenly speed up and they might not be able to stop it in time to prevent a crash. There was also uncertainty about the outcome, as the exact number of points lost could not be fully predicted from the marble stopping position. In fact, the bar was divided into 4 different payoff sections of equal length (60-46 points at the top; 45-26 and 25-16 points in the middle; 15-1 points at the end). If the marble crashed, 70-99 points would be lost. Within each section, the number of points lost was varied randomly from trial to trial.

At the beginning of “Together” trials, participants saw their own avatar next to the avatar of their co-player, and the marble in these trials was coloured green. Participants were instructed that, in these trials, both players would be playing together and either could use their mouse button to stop the marble. If neither player acted, the marble would crash and both players would lose the same number of points. If the co-player stopped the marble, the participant would not lose any points. If the participant stopped the marble, they would lose a number of points according to the position where they stopped it, and their co-player would not lose any points.

In fact, participants were playing alone in all trials, and the co-player's behaviour was simulated by the computer. The co-player's behaviour was programmed such that participants had to stop the marble in the majority of "Together" trials, to ensure a sufficient number of artefact-free trials was available for ERP analyses. If participants had stopped the marble more often than their co-player, and if participants did not act sooner, the co-player could stop the marble along the lower half of the bar. In that case, the marble would stop on its own, and participants received feedback of losing zero points. To avoid ambiguity about who caused the outcome, simultaneous actions of both participant and co-player were attributed to the participant. Thus, if the participant acted within 50ms of a simulated co-player action, this would count as participant's action, and feedback would indicate a loss according to the stop position.

### **6.3. ERP Preprocessing**

EEG-signals were processed using the Matlab-based open-source toolbox eeglab (Delorme & Makeig, 2004) with the ERPLab plugin (Lopez-Calderon & Luck, 2014).

The continuous EEG signal was notch-filtered and re-referenced to the averaged signal of the left and right mastoids. The signal was then cut into 3000ms epochs time-locked to the presentation of the outcome. Independent component analysis was used to remove eye movement artefacts. A 0.5 Hz highpass filter and a 20 Hz lowpass filter were applied. Epochs with signal artefacts were removed using an 80 $\mu$ V threshold. EEG signals were then averaged into event-related potentials separately for the two experimental conditions, using a 100ms pre-stimulus baseline. This resulted in an average of 40.78 (SD = 4.78) trials for the Alone condition (min = 34), and an average of 34.24 (SD = 8.41) trials for the Together condition (min = 21).

The FRN component was analysed as the mean amplitude between 250-330 ms, at electrode FCz, based on previous studies (Li et al., 2011; Yeung et al., 2005) and observation of grand ERPs and scalp topography.

## 6.4. Data analysis

Behavioural data in successful stopping trials (stopping position, outcomes, and agency ratings) and mean FRN amplitude were analysed using hierarchical linear regression models (i.e. linear mixed-effects models). This approach is advisable with unbalanced data, and allowed us to model single trial data (Baayen et al., 2008; Bagiella et al., 2000; Tibon & Levy, 2015). Models included the condition as a predictor, coded as Alone = 0, Together = 1. Where relevant, Stopping Position and Outcome were also included as covariates, after standardising the values within participants. All fixed effects were also modelled as participant random effects (random intercepts and slopes). Analyses were conducted using the *lme4* package (Bates et al., 2014) in R (R Core Team, 2015). Parameter estimates ( $b$ ) and their associated t-tests ( $t$ ,  $p$ ), calculated using the Satterthwaite approximation for degrees of freedom (Kuznetsova et al., 2015), are presented to show the magnitude of the effects, with bootstrapped 95% confidence intervals (Efron & Tibshirani, 1994).

Moreover, we analysed behavioural data (proportion of trials, agency ratings, and mean outcomes) from trials in which the marble crashed. ERP data for these trials was not analysed, however, due to low trial numbers. Finally, for together trials only, we compared the proportion of trials in which the co-player acted, relative to the marble crashing.

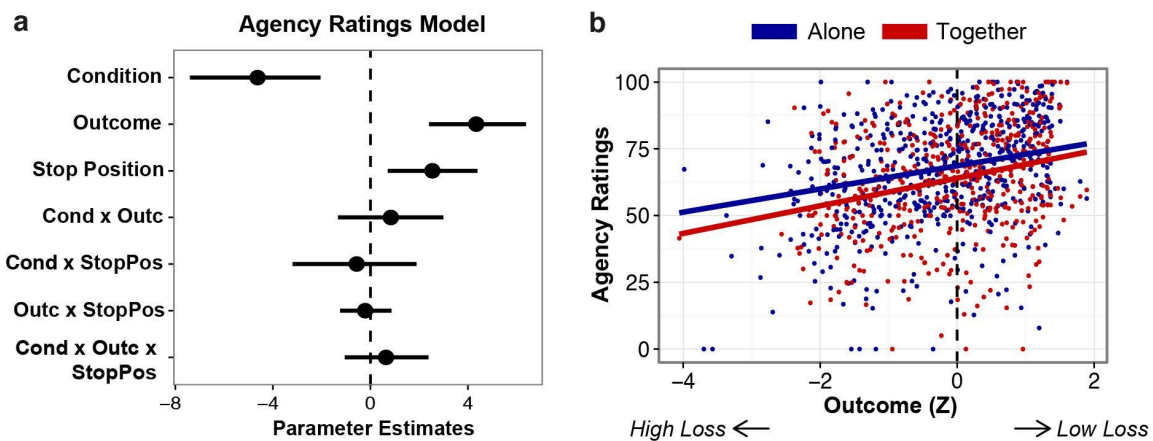
## 6.5. Results

### 6.5.1. Behaviour

The main focus of our analyses was trials in which the participant successfully stopped the marble. These trials were the same across the two social context conditions, but differed only in that participants acted while knowing that their co-player *could have acted instead* of them, in the together condition. To assess how participants' behaviour varied across social contexts, we modelled the position at which the marble was stopped. Participants stopped the marble significantly later in the together condition, relative to playing alone ( $b = 3.08$ ,  $t_{(554.04)} = 5.49$ ,  $p < 0.001$ ,

95% CI = [2.01, 4.15]). This suggests that participants waited longer to act in the together condition to allow time for their co-player to act instead of them.

Outcome (number of points lost) was predicted from the social context factor, stop position covariate, and their interaction. Outcomes were related to the marble stop position ( $b = 6.20$ ,  $t_{(28.89)} = 22.01$ ,  $p < 0.001$ , 95% CI = [5.59, 6.74]), with later stops resulting in smaller losses, as expected based on the task design. The social context did not influence outcomes ( $b = 0.089$ ,  $t_{(28.86)} = 0.27$ ,  $p = 0.79$ , 95% CI = [-0.63, 0.70]), nor did the social context by stop position interaction ( $b = -0.42$ ,  $t_{(61.51)} = -1.39$ ,  $p = 0.17$ , 95% CI = [-1.01, 0.23]). This shows that outcomes were similar across social contexts, for trials in which the participant successfully stopped the marble.



**Figure 6.2.** Behavioural results. **a** Parameter estimates for the model predicting agency ratings, with 95% bootstrapped confidence intervals. *Condition* refers to the effect of social context (Alone = 0 vs. Together = 1), such that a negative parameter estimate denotes a loss of agency in the Together condition. **b** Model predictions (lines) and mean agency ratings (dots) for the two social contexts as a function of outcome (number of points lost, Z-score). This displays the main effect of social context, with higher agency ratings in the alone condition. It additionally shows that agency ratings were lower for higher losses, and this relationship was similar across social contexts (no significant social context (condition) x outcome interaction, as seen in panel **a**).

Finally, agency ratings were modelled using social context, stop position, and outcome, plus their interactions. Results showed a significant reduction in agency ratings when playing together, relative to playing alone ( $b = -4.63$ ,  $t_{(21.59)} = -3.35$ ,

$p = 0.003$ , 95% CI = [-7.40, -2.05]; see **Figure 6.2.a**). Agency ratings were also predicted by the outcome ( $b = 4.34$ ,  $t_{(23.78)} = 4.76$ ,  $p < 0.001$ , 95% CI = [2.41, 6.38]), with smaller losses being associated with higher ratings. Finally, agency ratings were significantly influenced by the marble stopping position ( $b = 2.54$ ,  $t_{(21.17)} = 2.82$ ,  $p = 0.010$ , 95% CI = [0.71, 4.39]), with later stops being linked to higher ratings. There were no significant interactions (see **Figure 6.2.b**, and **Table 6.1** for full results).

**Table 6.1.** Agency ratings model: parameter estimates, with bootstrapped 95% confidence intervals. \* Based on Satterthwaite degrees of freedom estimation.

	Estimate	S.E.	$t$	$df^*$	$p^*$	95% C.I.	
						Lower	Upper
(Intercept)	68.65	3.07	22.36	24.01	< .001	61.98	74.65
Condition	-4.63	1.38	-3.35	21.59	0.003	-7.40	-2.05
Outcome	4.34	0.91	4.76	23.78	< .001	2.41	6.38
Stop Position	2.54	0.90	2.82	21.17	0.010	0.71	4.39
Condition x Outcome	0.84	1.18	0.71	22.04	0.487	-1.33	2.99
Condition x Stop Pos	-0.56	1.34	-0.42	23.96	0.678	-3.19	1.89
Outcome x Stop Pos	-0.21	0.56	-0.38	29.23	0.705	-1.24	0.86
Condition x Outcome x Stop Pos	0.64	0.83	0.77	25.15	0.450	-1.05	2.38

To check whether participants might have always reported less control in the together condition, agency ratings were analysed specifically in trials in which the marble crashed. Agency ratings were modelled by the social context, the outcome, and their interaction. When the marble crashed, results showed that only the outcome – how many points were lost – influenced agency ratings ( $b = 2.78$ ,  $t_{(23.69)} = 2.65$ ,  $p = 0.014$ , 95% CI = [0.73, 4.55]), with higher ratings associated with smaller losses. Social context no longer predicted agency ratings ( $b = 0.54$ ,  $t_{(33)} = 0.35$ ,  $p = 0.73$ , 95% CI = [-2.41, 3.42]), and there was no significant social context by outcome interaction ( $b = -0.37$ ,  $t_{(21.49)} = -0.22$ ,  $p = 0.83$ , 95% CI = [-3.79,

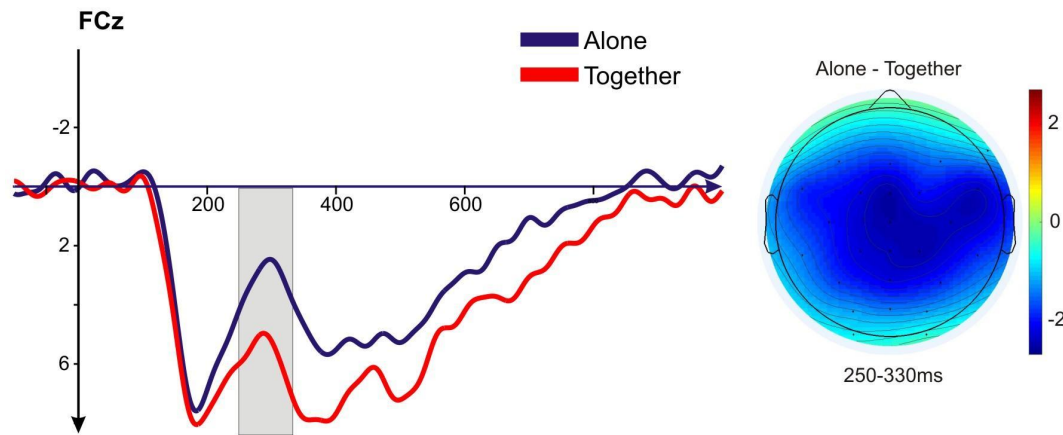
2.76]). We further checked that according to the task design, outcomes did not differ, on average, across social contexts (alone: mean = -84.92, SD = 2.88; together: mean = -85.34, SD = 3.41; paired samples t-test:  $t_{(24)} = 0.43$ ,  $p = 0.67$ ). Therefore, the relation between agency ratings and social context described above was specifically related to those trials in which the participant successfully acted.

To fully characterise participants' behaviour in the task, we also analysed number of trials in which the marble crashed, and in which the "Other" agent acted instead (in the together condition). The marble crashed significantly more often in the alone condition (mean = 19.47%, SD = 7.99), than when playing together (mean = 14.47%, SD = 6.08; paired samples t-test:  $t_{(24)} = 3.32$ ,  $p = 0.003$ ). In the together condition, the co-player acted significantly more often (mean = 19.33%, SD = 7.71) than the marble crashed (paired samples t-test:  $t_{(24)} = 4.60$ ,  $p < 0.001$ ). These results, together with the earlier finding of later stops in the together condition, show that participants adapted their behaviour in order to minimise their losses in the together condition, when the "co-player" could act instead of the participant. To assess whether this strategy really was beneficial, we averaged the outcomes across all trials (successful stops, marble crashes, and "co-player" actions) for each participant. Results confirmed that, overall, participants lost significantly less points in the together condition (mean = -20.86, SD = 3.78), relative to playing alone (mean = -27.71, SD = 4.43; paired samples t-test:  $t_{(24)} = -6.31$ ,  $p < 0.001$ ). Since the comparisons above showed no significant differences in outcomes across social contexts for successful stops, nor for marble crashes, this overall reduction in losses was clearly driven by the "co-player" action trials, in which the participant did not lose any points.

### 6.5.2. ERPs

Mean amplitudes for the FRN component were analysed with the same model as agency ratings. Results revealed that FRN amplitude was significantly reduced (i.e. more positive) when playing together, relative to the alone condition ( $b = 1.23$ ,  $t_{(161.77)} = 2.31$ ,  $p = 0.022$ , 95% CI = [0.15, 2.29]; see **Figure 6.3**). FRN amplitude was not significantly influenced by the outcome ( $b = 0.21$ ,  $t_{(43.11)} = 0.42$ ,  $p = 0.68$ , 95%

CI = [-0.74, 1.07]), nor by stop position ( $b = -0.62$ ,  $t_{(26.15)} = -1.16$ ,  $p = 0.26$ , 95% CI = [-1.63, 0.52]). There were no significant interactions (see **Table 6.2**).



**Figure 6.3.** Outcome-locked ERPs. Grand average time courses are shown for the two experimental conditions. The time window for the FRN (250-330ms) analysed is highlighted in grey. The topographic plot shows the scalp distribution of the difference between the social context conditions averaged across the FRN time window.

**Table 6.2.** FRN model: parameter estimates , with bootstrapped 95% confidence intervals.

\* Based on Satterthwaite degrees of freedom estimation.

	Estimate	S.E.	$t$	$df^*$	$p^*$	95% C.I.	
						Lower	Upper
(Intercept)	1.98	0.88	2.26	23.87	0.033	0.25	3.59
Condition	1.23	0.53	2.31	161.77	0.022	0.15	2.29
Outcome	0.21	0.50	0.42	43.11	0.68	-0.74	1.07
Stop Position	-0.62	0.53	-1.16	26.15	0.26	-1.63	0.52
Condition x Outcome	1.01	0.78	1.29	33.24	0.21	-0.49	2.65
Condition x Stop Pos	0.10	0.74	0.13	34.69	0.90	-1.36	1.63
Outcome x Stop Pos	-0.12	0.28	-0.45	42.52	0.66	-0.71	0.47
Condition x Outcome x Stop Pos	0.73	0.64	1.15	17.76	0.27	-0.59	1.99



## **6.6. Discussion**

To investigate the cognitive and neural consequences of the diffusion of responsibility phenomenon, we developed a task in which a costly action was required to prevent an even greater cost. Participants either played alone, or together with another agent who could act instead of them, sparing the participant any losses. The presence of a potential alternative agent led participants to act later, reduced their subjective sense of agency, and attenuated the neural processing of the consequences of their action, as reflected in reduced FRN amplitude.

### **6.6.1. Behaviour**

In trials in which the other player could act as well, participants acted later and rated their feeling of control over the action outcomes as lower. Importantly, the presence of the other player was manipulated on a trial-by-trial basis. Thus, our results show that behavioural decisions and sense of agency over action outcomes are continuously and flexibly updated in regard to the current social context.

When participants acted, they objectively had the same amount of control over the number of points they lost in “Alone” and “Together” trials. As soon as participants pressed the button, the marble stopped, indicating that they had successfully acted. Outcomes were always contingent on where the marble was stopped, even though they remained partially unpredictable. In contrast, when the alleged co-player acted, which was actually the computer, the marble stopped without a concurrent action from the participant. Thus, the loss in control reflected in the self-report was purely a subjective one, with objective action-outcome contingencies remaining constant.

Our results also showed that agency ratings were lower following greater losses. This is consistent with previous studies which also found reductions in the sense of agency for negative, relative to positive, outcomes, while using implicit measures of agency linked to the temporal perception of actions and outcomes (Takahata et al., 2012; Yoshie & Haggard, 2013). This presumably reflects the well-known self-serving bias of attributing negative outcomes to external factors, and positive outcomes to one’s own actions (Bandura, 1999). Importantly, this effect was not

related to, nor facilitated by, the presence of an alternative agent. That is, participants did not strategically attribute “away” their responsibility over negative outcomes to the other player. Interestingly, even when the marble crashed, which led to the worst outcomes, agency ratings were not influenced by social context but only by the actual outcome. Diffusion of responsibility therefore seems to occur beyond a self-serving bias of (mis)attributing negative outcomes to others.

Finally, higher agency ratings were associated with stopping the marble later, independently of the outcome. This effect could be due to a better match between the action performed and the theoretically optimal action in later stops. Alternatively, trials with early stops may be associated with more counterfactual thinking and regret over possibly smaller losses with later stops, linked to the self-serving bias mentioned above. Finally, outcomes of risky but successful choices could be perceived as more rewarding than those of less risky choices.

### **6.6.2. FRN**

ERP results showed an effect of social context on the neural processing of action outcomes. In otherwise identical trials, FRN amplitude in response to outcomes of successful actions was reduced by the presence of a co-player.

The FRN has been shown to be sensitive to the distribution of control among multiple players (Li et al., 2010), as well as to whether participants were instructed outcomes were random or controllable (Li et al., 2011; Yeung et al., 2005). We show a similar effect of social context on FRN amplitude, but now using a task in which instructions about action-outcome contingencies were identical between conditions, and full control over the outcome remained with the participant. Thus, the mere presence of another player was sufficient to evoke changes in the neural processing of action outcomes akin to those observed when control over an outcome is abolished or truly shared between multiple co-players. As such, our findings at the neural level offer an objective measure consistent with the perceived loss of control participants reported subjectively.

The FRN has been widely implicated in outcome monitoring (San Martín, 2012), and is thought to be sensitive to the motivational significance of outcomes (Holroyd & Yeung, 2012). As outcomes were perceived as less controllable and possibly less self-relevant in “Together” trials, the motivation to learn from such outcomes could be weakened, leading to a reduction in outcome monitoring.

### **6.6.3. Implications for concepts of diffusion of responsibility**

Both our findings of altered subjective sense of agency ratings and reduced FRN amplitude suggest that the co-player’s presence made participants feel less in control over the outcomes of their actions. This significantly extends current models of diffusion of responsibility (Bandura, 1999), by demonstrating an online effect of social context on the processing of action outcomes. This is in line with Bandura’s proposition that negative consequences of one’s own behaviour are less relevant to the self in a group than in an individual context (Bandura, 1999). We suggest that the diffusion of responsibility effect may at least partially be due to people actually experiencing the consequences of their actions as less controllable and thus less meaningful in social contexts, relative to when acting alone.

The mechanisms by which the presence of others leads to a reduced coupling of outcomes to the actions that caused them remain unclear. Different aspects of the social situation could potentially affect people’s sense of agency in the absence of an objective reduction of control. Cueing the ‘together’ condition possibly triggered counterfactual thinking about what would have happened if the participant had chosen not to act. The presence of another player presumably made alternative scenarios, i.e. in which not acting could have led to better (no-loss) outcomes, more accessible and possibly also reduced participants’ motivation to act.

Such counterfactual thinking might take place once the outcome is known, reflecting a retrospective influence on outcome monitoring, and sense of agency. However, it might also influence the decision making in advance of the action. Having to consider the potential behaviour of an alternative agent would increase the complexity of decision making, by adding extra informational load. Previous research has shown

that difficulty of action selection can reduce sense of agency (Chambon, Sidarus, et al., 2014). Importantly, such processes are thought to influence the sense of agency *prospectively*, as they take place long before the outcome is known. From this perspective, the increased complexity in decision making in social situations could partly explain how the presence of another agent might affect behaviour, reduce outcome monitoring and reduce the sense of agency.

Interestingly, Bandura (1991) distinguished *diffusion* of responsibility, linked to the presence of other agents, from situations in which one's responsibility is *displaced* onto perceived authority figures guiding one's actions. Consistent with our findings, a recent study showed that coercion reduces sense of agency and attenuates the sensory processing of action outcomes (whether neutral or negative), relative to free choice (Caspar et al., 2016). Aside from social contexts, Bandura (1991) further proposed that one's sense of responsibility might be reduced by disregarding or distorting action consequences, i.e. inattention or a misrepresentation of outcomes. While his model distinguishes these three components, the findings discussed here suggest that changes in the processing of action outcomes may, in fact, underlie the effects of social context on the sense of agency and responsibility. Previous models of sense of agency emphasise that it involves the integration of multiple cues, ranging from sensorimotor, to affective, to contextual information (Synofzik et al., 2013). These results suggest that social context constitutes another influence on sense of agency, which may partly act through altering outcome monitoring, but might also have independent effects. For example, a situation like coercion involves a restriction of one's choice in addition to a social component, and freedom of choice is known to influence sense of agency independently of social context (Barlas & Obhi, 2013).

Our results also suggest possible mechanisms whereby diffusion of responsibility may causally affect people's behaviour. We deliberately designed our task to eliminate ambiguity in agency attribution, that is, about *who* caused a given outcome. However, in real-life social situations, attribution of outcomes is likely to be more ambiguous than in comparable non-social situations. If the neural processing and

monitoring of action consequences is then reduced by the presence of other potential agents, this may increase the likelihood of attribution errors. Thus, reduced outcome processing due to diffusion of responsibility may lead to ambiguity in attribution. Correctly attributing consequences to their causes, however, is a critical prerequisite for learning action-outcome associations and likely also for forming a sense of moral responsibility. In this sense, the social dilution of agency might potentially have both immediate and longer-term effects on agency learning.

### **6.6.4. Limitations and outlook**

Research on diffusion of responsibility has mostly focused on the possibility that social contexts reduce individuals' sense of responsibility for negative consequences. It remains to be shown whether the effects found in the current study generalize to neutral and positive action outcomes. Though caution is required in interpreting any null result, the absence of an interaction between outcome and social context in the present study suggests that this may be the case. Moreover, our task involved negative consequences to the self; therefore, effects could differ when the consequences of our actions are experienced by other people. Interestingly, in Caspar et al.'s (2016) study, coercion led to a reduction in the processing of action outcomes which were positive for oneself, but at the same time could be harmful to others.

In this study we primarily focused on the processing of outcomes following participants' actions. However, perceived responsibility, and sense of agency can differ for action relative to inaction (Baron & Ritov, 2004; Kordes-de Vaal, 1996). At a behavioural level, we found no evidence for diffusion of responsibility for negative outcomes following inaction. It remains to be shown whether there are systematic differences in social context effects on sense of agency related to action and inaction, and how these effects translate into behaviour. While it seems likely that people may feel less inclined to act in a situation that they perceive as less controllable, the interactions between sense of agency, action and inaction are yet to be investigated.

Finally, our conclusions regarding sense of agency are based on an explicit rating of control. While our findings are corroborated by the ERP results, future studies could additionally use implicit measures of agency to further elucidate the effects of social context on pre-reflective aspects of sense of agency.

### **6.6.5. Conclusions**

We show that diffusion of responsibility is not merely a post-hoc phenomenon reflecting a self-serving bias, but an online influence on how people process the consequences of their actions. The presence of other agents, even in the absence of attributional ambiguity, can lead to a reduction in outcome monitoring and a reduction in individual sense of agency. These effects could be mediated by counterfactual thinking or increased complexity of decision making processes in social situations.







## Chapter 7. General Discussion

The present thesis investigated how action selection processes contribute to the sense of agency (SoA). Metacognitive monitoring of action selection processes can inform the SoA of any difficulties, or disruptions, experienced during action selection. These signals are prospective, as they emerge at the time of the action, and long before the outcome is known. This was demonstrated here in the EEG study described in Chapter 4, and previously in fMRI (Chambon et al., 2013), TMS (Chambon, Moore, et al., 2014), and MEG (Stenner, Bauer, Heinze, et al., 2014).

This final section will summarise the novel insights offered by our research into the mechanisms underlying the prospective contributions to SoA. We have shown that the effects of action selection on SoA are robust, and generalise across a number of tasks, even to dynamic contexts. Moreover, our research highlights the integrative nature of SoA. We will briefly consider how SoA combines a variety of cues, whether prospective, retrospective, or contextual. Finally, we consider the potential implications of a prospective SoA, namely to our societal notions of responsibility. We also suggest that the complexity of decision making required in social contexts may mediate the well-known *diffusion of responsibility* effect.

## 7.1. Prospective Sense of Agency

### 7.1.1. Effects of Disrupted Action Selection

The studies presented throughout the thesis show that SoA is sensitive to disruptions in action selection. Various manipulations were used, targeting different stages of action selection: 1. Dysfluency in visual processing (Chapter 3, Exp. 1); 2. Ambiguity in stimulus categorisation (Chapter 3, Exp. 2); 3. Response conflict, using the Eriksen flanker task (Chapter 2, Chapter 3, Exp. 3), or subliminal priming (Chapter 4).

Some studies involved a simple chain of events: a target, an action, and an outcome (Chapters 2 & 4). Participants gave judgements of agency over action outcomes. These judgements were mostly collected at the end of each trial, but can also be made at the end of a block (half of the participants in Exp. 1 of Chapter 2; cf. Wenke et al., 2010). Other studies used a dynamic computer game (Chapter 3), in which many potential targets, actions, and outcomes were present. Participants gave judgements of agency over the game, after playing for 30 s. Therefore, disruptions to action selection can have online effects on SoA. Additionally, fluency/conflict experiences can also accumulate over time, and even become associated with particular outcomes.

Our experiments also support a particular role of dysfluency *in action selection*, rather than a more global fluency effect, e.g. in stimulus processing. Even though neutral flankers were perceptually dissimilar from the target letter, this did not lead to a significant reduction in agency ratings relative to congruent flankers in Experiment 1 of Chapter 2. Moreover, flankers that appeared 100ms after the target still affected action selection and the SoA (Experiment 3, Chapter 2), which would be inconsistent with an account in terms of “conceptual congruency” between an initial intention, e.g. triggered by a subliminal prime, and the executed action (cf. Chambon & Haggard, 2012). Chapter 3 did reveal that SoA is sensitive to disruptions of fluency in other aspects of cognitive processing besides response conflict, such as dysfluent stimulus processing. Yet, we suggest that the effects seen here are specifically

related to action selection and agency, since the manipulations of fluency regarded processes that subserved action selection processes.

The literature on fluency effect has shown a myriad effects on various judgements (Alter & Oppenheimer, 2009). Further research is needed to clarify whether similar or distinct mechanisms are involved across different metacognitive domains, such as confidence judgements, judgements of learning, or the SoA. One distinguishing feature of SoA is that it involves integrating information about both one's action and about the outcome(s) of that action, in order to establish their relation (see also section 7.5.2). This was evidenced in Chapter 4, as agency ratings were correlated to both neural correlates of action and of outcome monitoring. For example, confidence judgements about the accuracy of a response would presumably only be concerned with information about the action. Therefore, while there may be overlap in some metacognitive processes that may inform different judgements, differences may also arise due to the different sources of information that are relevant for each particular judgement. Notably, the fluency literature has shown that when people are aware that the source of fluency experiences is unrelated to the task at hand, they are able to discount it (cf. Alter & Oppenheimer, 2009). Chapter 2 showed that even when the source of dysfluency in action selection was made evident – by using supraliminal flankers – fluency effects on SoA remained. This might arguably reflect that such dysfluency was considered particularly relevant for the SoA. Yet, it remains to be shown whether truly task-unrelated fluency/dysfluency experiences could still influence SoA.

Earlier studies suggested that fluency in action selection might *enhance* SoA (Wenke et al., 2010). However, the results observed could be due to fluency enhancing SoA, or dysfluency leading to a reduction in SoA. Our work suggests that it is more likely that dysfluency, or conflict, in action selection leads to a reduction in SoA. In Chapter 2, Experiment 1 showed that although both facilitation and conflict effects were found for action selection, only conflict affected SoA significantly (see **Figure 2.2**, p. 58). Experiment 3 also showed that conflict led to a reduction in SoA, even when motor performance was not affected (see **Figure 2.4**, p. 70; cf. Chambon & Haggard,

2012). In Chapter 4, we found that SoA was associated with an ERP component, the correct-related negativity (CRN), which is thought to reflect conflict monitoring (Larson et al., 2014).

While fluent processing can be associated with a positive affect (Reber, Winkielman, & Schwarz, 1998), conflict signals are especially motivationally significant because they may indicate a need to adjust subsequent behaviour (Botvinick & Braver, 2015; Holroyd & Yeung, 2012). Moreover, a strong SoA may be a default state, requiring evidence to the contrary (Sidarus et al., 2013). Therefore, conflict, or dysfluency, signals can have a greater impact on SoA, than signals indicating fluent selection. Similarly, outcome monitoring is associated with “error” signals, indicating a mismatch between intended/predicted and actual action outcomes (Blakemore et al., 2002). The angular gyrus (AG), in the inferior parietal lobe (IPL), has been implicated in the online monitoring of disruptions both to action selection (Chambon et al., 2013) and to outcome monitoring (Farrer, Frey, et al., 2008), with greater activity linked to a subjective loss of agency. The SoA is especially tuned to detecting disruptions in the smooth flow of voluntary action, from intention, to action, to outcome.

### 7.1.2. SoA and Post-decisional Action Monitoring

Our results further suggest that prospective cues to SoA are related to post-decisional metacognitive signals. In the EEG study reported in Chapter 4, we found that agency ratings were related to the action-locked CRN component (see **Figure 4.8.a**, p. 134). This has been previously linked to conflict and error monitoring (Larson et al., 2014). As this component arises at the time of action execution, it reflects *post*-decisional processes. Such processes are linked to evidence accumulation about the correct, or adequate, response, which can continue after an initial decision, but may precede or be concurrent with the action itself (Yeung & Summerfield, 2014). We also found that *pre*-decisional monitoring was sensitive to response conflict, as indexed by the target-locked N2 potential (see **Figure 4.3**, p. 127). However, N2 was not linked to agency ratings. Consistently, a role for post-decisional processing had been previous shown in a TMS study. Disrupting IPL immediately after the action abolished selection fluency effects to SoA, similarly to

when applied during action selection (Chambon, Moore, et al., 2014). Post-decisional processes can integrate both early and late stages of action selection, thus providing a more complete signal to SoA.

Interestingly, post-decisional processes related to action selection have also been implicated in models of confidence judgements (Fleming et al., 2014; for a review, see Yeung & Summerfield, 2014). Namely, the CRN component has been associated with confidence judgements (Boldt & Yeung, 2015; Scheffers & Coles, 2000). Moreover, studies on confidence often employ manipulations of processing fluency, categorisation ambiguity, or uncertainty, similarly to the manipulations we used in Chapter 3, which we found affect agency judgements. This suggests that confidence and agency judgements may rely on common metacognitive signals.

We may speculate that, in the context of agency, confidence that the action is appropriate (“I just know what to do”) may be related to confidence that the action outcome will be as intended. Alternatively, low-confidence actions may be associated with uncertainty about action consequences. Sometimes, we may not know what outcomes to expect, and thus a sense of confidence in our actions (“I did the best I could”) may be more important to SoA than the consequences.

Similar metacognitive signals may be associated with other aspects of processing quality, namely processing dysfluency, uncertainty, response conflict. Such signals may be vague, and unspecific regarding their source, and thus have global influence on other processes (Winkielman et al., 2015). Therefore, selection fluency may be used as an umbrella term, which encompasses various disruptions to action selection, such as conflict, uncertainty, or difficulty.

Fluency should not be confused with speed of processing. This may often be related to fluency, or may be used as a proxy for difficulty. For example, both evidence and decision time have been shown to influence confidence judgements (Kiani et al., 2014). Indeed, we found that RTs were negatively related to agency ratings in Chapter 4, but selection fluency (i.e. congruency) still had a significant, independent effect on agency (see **Figure 4.7**, p. 133; see also **Appendix A** for similar results

from Chapter 2). Moreover, RTs can also be dissociated from the effects of conflict on SoA. In Chapter 2, Experiment 3 showed that when flankers preceded the target by 500 ms, congruency effects on motor performance were abolished, yet incongruent flankers still led to lower SoA than congruent flankers (see **Figure 2.4**, p. 70). Even when RTs are faster for incongruent, relative to congruent, subliminal priming, SoA is still reduced by incongruent priming (Chambon & Haggard, 2012). Therefore, RT monitoring may serve as an additional, independent cue to agency and confidence judgements.

Finally, in Chapter 4, we suggested that SoA is linked to a post-decisional signal that indicates unresolved conflict at the time of action, rather than to initial signals linked to conflict detection or cognitive control recruitment. This may appear inconsistent with the suggestion in Chapter 2 that SoA is especially sensitive to detecting the presence of conflicting intentions, even when conflict is eventually resolved. However, we note that in Chapter 4 we found a general relation between CRN and agency ratings, using a trial-wise analysis. On the other hand, the study in Chapter 2 involved comparing average agency ratings across congruency conditions. Whilst incongruent trials will more likely trigger conflict, conflict signals can also vary independently of congruency. Additionally, N2 amplitude was also linked to cognitive control recruitment, rather than just conflict detection, as congruency effects were reduced for slow RTs (see **Figure 4.4**, p. 128). Therefore, trials with a large N2 amplitude may reflect instances in which successful cognitive control prevented or limited response conflict. In those cases, SoA might be preserved, rather than reduced.

## 7.2. Integrating Prospective and Other Cues to Sense of Agency

In line with previous proposals (Moore & Fletcher, 2012; Synofzik et al., 2013), our work shows that the SoA involves integration of a variety of cues. These cues may be of different types, e.g. sensorimotor, or conceptual, and may become available at different times, e.g. at the time of the action, or of the outcome. Critically, the weighting of these cues, and whether they have independent or interrelated effects

on SoA is dynamically updated, and highly dependent on context and availability of cues. SoA is highly flexible and adaptable to current context and task demands.

### **7.2.1. Role of Awareness of Biases and Choice**

Our work shows that the effects of action selection on SoA are independent of whether one is aware, or not, of the stimuli that influence action selection. We showed a consistent reduction in SoA due to disruptions to action selection induced subliminally (Chapter 4), and supraliminally (Chapters 2 & 3). Previous studies used subliminal priming of actions to ensure that any effects found could not be attributed to participants knowing that their actions were manipulated. Moreover, it was unclear whether selection fluency might interact with awareness. We demonstrate that unawareness of biases is not necessary to investigate the effects of action selection to SoA. [But see Damen et al. (2014), and section 7.3 below.]

The effects of selection fluency on agency were also independent of choosing freely between action alternatives, or following instructions (Chapter 4 & Experiment 2 in Chapter 2; cf. Wenke et al., 2010). Moreover, we found that free choice trials were associated with a higher SoA than forced choice, in the subliminal priming study (Chapter 4). However, there was no robust effect of choice when using supraliminal flankers (Chapter 2, Experiment 2). The combination of free and forced choice trials, combined with awareness of the flankers might have led participants to never feel very free, even when they could choose between a left or right key press. Freely choosing our actions, relative to following instructions, is thought to be central to our SoA. Indeed, it has been linked to SoA (Barlas & Obhi, 2013; Caspar et al., 2016; Wenke et al., 2010). However, it has been suggested that our sense of freedom is only one aspect of SoA (Pacherie, 2008). Therefore, it might serve as another independent cue to SoA.

### **7.2.2. Lack of Relation to Outcome Predictions**

The EEG study in Chapter 4 further showed that the influence of selection fluency on SoA is not dependent on changes in outcome processing. We found that outcome monitoring, indexed by the outcome-locked feedback-related negativity (FRN)

component, was associated with agency ratings (**Figure 4.8.b**, p. 134). However, we did not find any direct effects of selection fluency on this measure of outcome processing. Additionally, we found no relation between the action-locked CRN and the outcome-locked FRN, even though both were associated with agency ratings. Therefore, the CRN and FRN components reflect two independent processes that, respectively, make prospective and retrospective contributions to SoA (see **Figure 4.9**, p. 136).

In line with this, the multi-study analysis described in Chapter 5 showed that selection fluency can have a general effect on SoA throughout instrumental learning, when participants rely on action-outcome interval to guide their agency judgements (experiments in Group 2, see **Figure 5.5**, p. 159). This general effect of selection fluency suggests that we learn to use it as a heuristic cue to SoA. Typically, dysfluent selection will be predictive of unsuccessful or unexpected outcomes. These effects may also be associated with, or the heuristics mediated by, the affective consequences of conflict, or dysfluency. This would be in line with other fluency effects on judgements (Alter & Oppenheimer, 2009; Winkielman et al., 2015).

### 7.2.3. Interactions with Outcome Monitoring

When faced with a new environment, an optimal cue integration account might have predicted that the poor reliability of action-outcome knowledge in a new environment could be compensated by a greater influence of selection fluency to SoA. If so, as reliable action-outcome contingencies were learnt, less reliable cues based on selection fluency should have decreasing effects. Our multi-study analysis (Chapter 5), in fact, showed that any interactions between instrumental learning and selection fluency were in the opposite direction. That is, the effects of selection fluency on SoA *increased* during instrumental learning (see **Figure 5.5**, p. 159). Interestingly, this occurred only when participants were focused on learning action-outcome contingencies (the experiments from Group 1, described in Chapters 2 and 4).

We ruled out the possibility that people learned, in a context-specific manner, to use action selection processes as a proxy for true action-outcome contingency. This



account would not be consistent with the aforementioned finding of a general, heuristic effect of selection fluency on SoA, even in a new environment (Group 2 experiments). Instead, we propose the novel suggestion that this interaction arises because action selection processes directly influence instrumental learning rates. That is, conflict during action selection could disrupt the learning of action-outcome associations, which would result in a slower learning rate (in linking dysfluent action - outcome), compared to fluency (i.e. linking fluent action - outcome). This proposal would be compatible with a heuristic use of selection fluency as a cue to SoA, and may perhaps even underlie the acquisition of this heuristic. Yet, it remains speculative, since we only measured agency ratings, rather than directly test knowledge of action-outcome contingencies. Notwithstanding that limitation, we believe agency ratings could reflect outcome knowledge indirectly, since participants were instructed to attend to the outcomes, and we observed an increase in agency ratings across the trials (for Group 1 studies, in which the interaction was found).

This account could also explain how conflict experiences become associated with specific outcomes, as seen when collecting agency ratings at the end of a block (half of participants in Experiment 1, Chapter 2; Wenke et al., 2010). From this perspective, lower agency ratings for outcomes that follow dysfluent selection would be related to poorer knowledge about action-outcome contingency. Interestingly, once fully predictable action-outcome contingencies have been well-learned, subliminal priming of actions can influence sensory attenuation (Stenner, Bauer, Sidarus, et al., 2014), and sensory predictions in sensory regions, even before an action was made (Stenner, Bauer, Heinze, et al., 2014). This supports the idea that selection fluency can affect the linkage between action and outcome. Also worth noting, studies on metacognition have shown that high confidence errors, e.g. on general knowledge questions, are associated with better error correction, by learning from feedback, than low confidence errors (Metcalf, Butterfield, Habeck, & Stern, 2012). This suggests that the link between (erroneous) responses and their outcomes (corrective feedback) was weakened for low confidence responses.

More consistent with an optimal cue integration account, one study found that action selection had a stronger influence on SoA when outcomes were unexpected (Sidarus et al., 2013). This study used partial action-outcome contingency (67%), which meant there was some “expected uncertainty” (Yu & Dayan, 2005) regarding the outcomes. That is, the occasional violation of outcome predictions was expected, and not diagnostic to SoA. Therefore, this may have led to a reduction of the contribution of outcomes to SoA when predictions were violated, and a relative increase in the weighting of action selection processes on SoA. Moreover, results also showed that when action selection was fluent, outcome expectation had a reduced effect on SoA. This suggested a reciprocal influence between the two cues, depending on their signal, and not only their general reliability.

Those studies involved distal and abstract action-outcome contingencies. Chapter 3 used a videogame-like task to manipulate action selection, and additionally manipulated the proximal outcomes of action, i.e. the means to achieve a goal. Adding a discrepancy between participants’ mouse movements and the movements of a cursor on the screen, led to a large reduction in SoA, which could overshadow the effects of action selection (Exps. 1 & 2). These findings are also consistent with optimal cue integration, as accurate motor control would be a highly reliable and salient indicator of agency over the game.

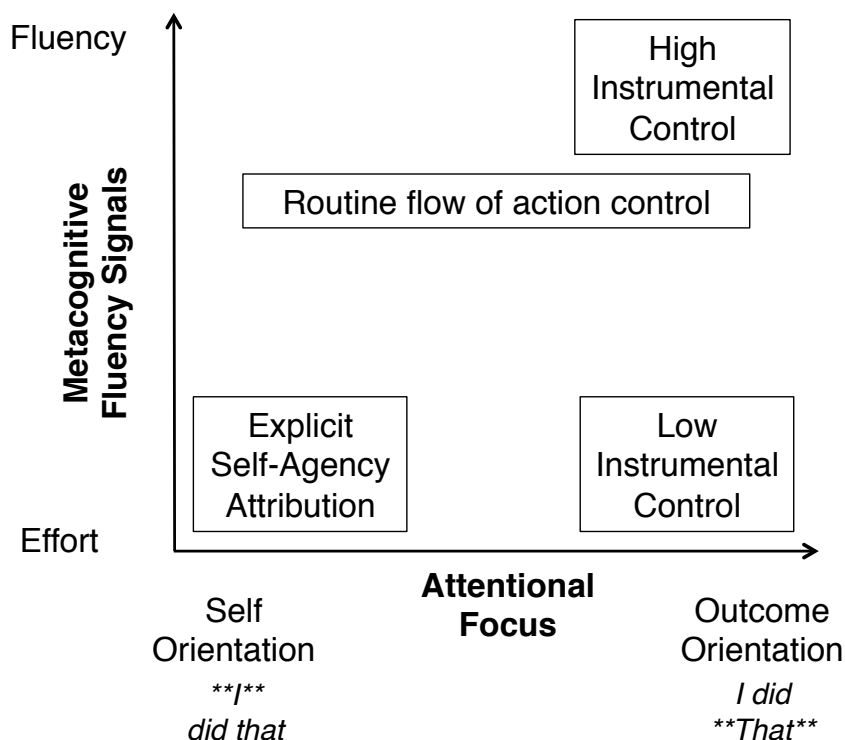
Therefore, the type of outcomes (proximal means vs. distal ends), and whether action-outcome associations are still being learnt, or are already known, may impact on the effects of action selection on SoA. Additionally, knowledge about the outcomes may affect the mechanisms through which action selection influences SoA, that is, whether it changes: the learning process itself; outcome predictions; or the weighting of outcome information.

### **7.3. Varieties of Agency Experience**

The research discussed in the present thesis demonstrates the flexibility and complexity of our experience of agency. We also believe that there are several *varieties* of agency experience, as mentioned in the introduction (section 1.4.1, p.

41). Previous work has often limited SoA to agency attribution. In contrast, our work examined SoA in terms of our experience of instrumental control. Moreover, while our work has shown that selection dysfluency reduces SoA, others have proposed that effort can increase SoA (Damen et al., 2014; Demanet, De Baene, et al., 2013; Lafargue & Franck, 2009). We now tentatively propose a model that explains distinctive varieties of agency experience based on two independent axes: attentional focus, and metacognitive signals (see **Figure 7.1**).

On the one hand, agency involves self-other distinction. When our attention is drawn towards the self, metacognitive signals of effort can highlight the sense that *\*\*I\*\** caused a given outcome, that is, the experience of explicit self-agency attribution (“*I* did that”). This may be especially relevant when action-outcome contingencies are well known – violations of expectations may be due to other’s actions, rather than real contingency changes.



**Figure 7.1.** Four distinctive varieties of agency experience can be explained by independent axes related to focus of attention, and to metacognitive signals. All four varieties of agency experience are assumed to depend on some degree of action-outcome contingency, and are all assumed to contribute to explicit measures of the Sense of Agency.

On the other hand, agency also involves linking our actions to specific outcomes. When we are focused on tracking the outcomes of our actions, metacognitive signals of fluency will boost our experience of instrumental control (“I did *that*”). However, if our actions are dysfluent, or effortful, our sense of instrumental control will be low. This will be particularly relevant when faced with a new environment, in which we will be focused on understanding the instrumental relation between our actions and external events. During instrumental learning, agency attribution may be less relevant to SoA.

Finally, the routine flow of everyday action control will normally be associated with metacognitive signals of fluency, as we will be familiar with the necessary actions and consequences. This experience may span across instances of greater focus on the self, or on the outcomes of action.

#### **7.4. Implications for Responsibility**

Previous work on decision making and social psychology has suggested that perceived responsibility, for our actions and those of others, in addition to behaviour, can be influenced by decision making processes. Difficulty, or uncertainty, in decision making can lead people to avoid making decisions altogether (Iyengar & Lepper, 2000). Even when people do act, they may feel less responsible for the consequences of their actions, especially for negative outcomes (Heath & Tversky, 1991).

Notably, social contexts can dramatically increase the complexity of decision making processes, as well as of outcome monitoring. On the one hand, the potential actions of others, and the consequences of those actions, need to be considered in addition to our own potential actions and outcomes. On the other, ambiguity about authorship of outcomes is increased. We suggest that this complexity may mediate the well-known *diffusion of responsibility* effect, in which people report feeling less responsible for acting in the presence of others (Bandura, 1991).

In the EEG study described in Chapter 6, we developed a task in which we controlled for ambiguity about outcome attribution, but varied social context. The only difference between conditions was whether participants believed another agent was present, who could act instead of them, or not. Therefore, this could have a prospective effect on behaviour, outcome monitoring, and the sense of agency. Our results confirmed that social context led to changes in behaviour, as participants waited longer for the other agent to act. It additionally led to a reduction in outcome monitoring, indexed by FRN, and sense of agency, relative to playing alone. While these last two effects could be explained by counterfactual thinking after the action, a prospective effect seems to better explain changes in behaviour. Moreover, this study showed that diffusion of responsibility is not merely a post-hoc justification, linked to a self-serving bias, but is associated with online changes in the experience of agency.

Nonetheless, it is important to recognise that our SoA can be affected by processes that do not *directly* reflect the relation between our actions and their consequences. Prospective cues to SoA may influence our sense of responsibility. Selection fluency may be statistically related to predicted outcomes, and hence serve as a convenient proxy, however, it does not actually affect action-outcome contingencies. Being aware of this may allow us to more strategically use selection fluency as a cue to SoA, or responsibility, when relevant, but not be misled by it when irrelevant, e.g. competence in sports inflating perceived chances of winning a bet on a game (Heath & Tversky, 1991). Similarly, understanding that social context can lead to a reduction in our SoA independently of actual changes in agency, may allow us to take that into account when assessing our responsibility for acting, or over the consequences of our actions. This may help reduce the negative consequences of diffusion of responsibility.

## 7.5. Remaining Issues

### 7.5.1. Fluency vs. Effort

In Chapter 2, we investigated three possible factors that could influence whether SoA is enhanced or reduced by selection fluency vs. effort. We tested the effects of

awareness of biasing influences, of choice, and of the timing of conflicting stimuli. Yet, our results consistently showed that effort led to a reduction in SoA. We suggest this may be related to the fact that our experiments focus on the instrumental aspect of agency (see also section 7.3 above). Yet, it remains unclear whether the differential patterns of effects found here relative to other studies (e.g. Damen et al., 2014) may relate to differences in the agency measures (asking about control vs. attribution), or in salient cues to agency (variability in outcome identity vs. action-outcome interval). Clearly, further research is needed to clarify the conditions under which fluency vs. effort enhance SoA.

### **7.5.2. Control over Actions vs. Control over Outcomes**

It remains unclear whether the effects of selection fluency on SoA specifically reflect an unintentional conflation between fluency/conflict experiences and SoA, when the two should be clearly dissociated. Alternatively, people may be able to adequately dissect the components of controlling their actions vs. controlling action outcomes. Whereas objective control over actions is affected by selection fluency, selection fluency cannot influence action-outcome contingencies. Here, we asked people to judge their control over action outcomes, but we found reliable effects of the ease versus difficulty of selecting the action itself. It has been previously shown that response conflict leads to a reduction in SoA over actions (Morsella et al., 2009).

This distinction becomes especially relevant if we consider individual differences. For example, it will impact how we interpret the fact that a person, or group (e.g. patients), does not show fluency effects on SoA. If we simply ask people to judge SoA over action outcomes, we cannot distinguish whether the absent effect is due to a reduced sensitivity to metacognitive signals about action selection, or whether those metacognitive signals were not used in agency judgements, arguably reflecting someone who accurately followed instructions to judge action-outcome contingencies. Future work is needed to clarify how these two aspects of SoA relate to each other, and how they are related to the interaction between cues to SoA.

### 7.5.3. What is Prospective Agency for?

Finally, we suggest that prospective SoA may offer a mechanism through which we can intervene, more quickly, and effectively, in the course of events triggered by our actions. Difficulty or uncertainty experienced during decision making may serve as a signal that further steps may be necessary to ensure the predicted, or desirable, consequences will follow. Such prospective interventions seem precluded by retrospective accounts of SoA, since outcomes must be known in order to be compared to predictions.

Research on the metacognition of action similarly highlights a role for metacognitive signals about action selection for adjusting subsequent behaviour (Botvinick & Braver, 2015; Yeung & Summerfield, 2014). However, it remains unclear whether and how a prospective *experience* of agency may impact such behavioural adjustments.





## Appendix A. Supplementary Analyses for Chapter 2

These supplementary analyses aimed to assess whether the congruency effects on agency ratings observed with ANOVA might be explained through RT monitoring. For this, we used hierarchical linear regression models to predict agency ratings from the experimental factor Congruency, and included RTs as an additional predictor. RTs were standardised within-participants, and all effects were entered as fixed and a participant varying effects. Other experimental factors were excluded for simplicity, and since they did not significantly affect agency ratings. Moreover, only Experiments 2 & 3 were analysed, since half of participants in Experiment 1 only gave block-wise ratings.

**Table A1.** Model for Experiment 2 predicting agency ratings by experimental factors and RTs (Z, within-participants): parameter estimates, with bootstrapped 95% confidence intervals. \* Based on Satterthwaite degrees of freedom estimation.

	Estimate	SE	<i>t</i>	df *	<i>p</i> *	95% CI	
						Lower	Estimate
(Intercept)	6.025	0.227	26.52	23.0	0.000	0.000	5.580
<b>Congruency</b>	<b>0.207</b>	<b>0.061</b>	<b>3.39</b>	<b>23.0</b>	<b>0.003</b>	<b>0.003</b>	<b>0.087</b>
RT (Z)	-0.158	0.032	-4.98	22.9	0.000	0.000	-0.220

**Table A2.** Model for Experiment 3 predicting agency ratings by experimental factors and RTs (Z, within-participants): parameter estimates, with bootstrapped 95% confidence intervals. \* Based on Satterthwaite degrees of freedom estimation.

	Estimate	SE	<i>t</i>	df *	<i>p</i> *	95% CI	
						Lower	Estimate
(Intercept)	6.547	0.297	22.04	24.0	0.000	5.965	7.130
<b>Congruency</b>	<b>0.110</b>	<b>0.058</b>	<b>1.91</b>	<b>23.8</b>	<b>0.069</b>	<b>-0.003</b>	<b>0.224</b>
RT (Z)	-0.093	0.032	-2.91	24.0	0.008	-0.155	-0.030

Consistent with the ANOVA findings, Congruency was an important predictor of agency ratings in these models. Although the effect is only marginally significant for

Experiment 3, this was also the case in the ANOVA analyses, thus it reflects the weak congruency effect found in that experiment.

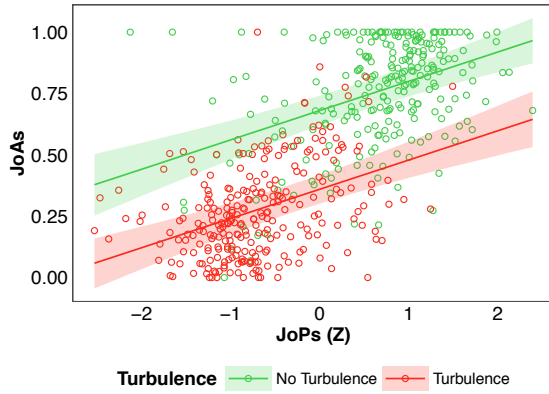
In line with the findings of Chapter 4, RTs were related to agency ratings, suggesting that RT monitoring may serve as an additional input to the sense of agency, but it cannot fully explain Congruency effects.

## Appendix B. Supplementary Analyses for Chapter 3

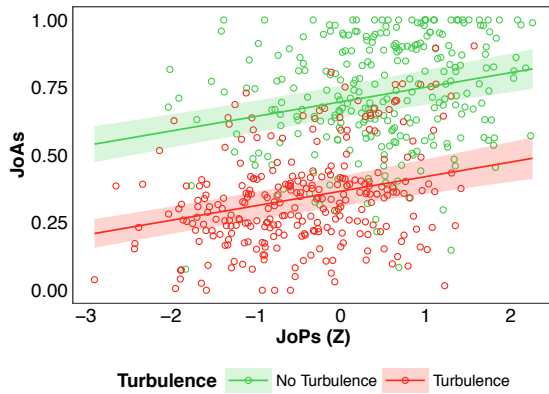
In our models of judgements of agency (JoAs), we used hierarchical linear regression to assess the effect of our experimental manipulations on JoAs, after accounting for a linear relation between JoAs and judgements of performance (JoPs). The assumption of a linear relation between JoAs and JoPs was motivated by the hypothesis that monitoring one's performance would have a general influence on JoAs. Non-linearities in this relation could hypothetically emerge if JoAs were especially related to JoPs for particular levels of performance (e.g. high vs. low), for example. Such potential non-linearities could in turn question any conclusions drawn about the effects on JoAs of other predictors in the model (i.e. our experimental factors).

To clarify this issue, we plotted the data across the three experiments, to observe the relation between JoAs and JoPs. Additionally, as the turbulence manipulation (i.e. discrepant visual feedback) had a very large effect on both JoPs and JoAs, this factor was considered as it might explain data clustering at higher vs. lower performance. The other factors used to manipulate action selection across experiments had quite small effects, thus they were not considered.

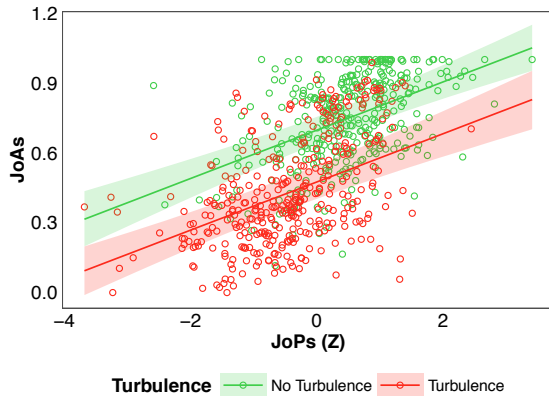
The plots below display mean JoAs (open circles) and model predictions (lines, with shaded 95% CIs) as a function of JoPs, and of the turbulence factor. Model predictions were obtained from 10000 simulations from the posterior distribution of plausible parameter values under uniform priors (Gelman & Su, 2015). Across experiments, these figures appear to support the assumption of a linear relation between JoAs and JoPs, and demonstrate the clear impact of turbulence on both variables.



**Figure B1.** Experiment 1: Mean JoAs (circles) and model predictions (regression line, and shaded 95% prediction intervals) as a function of JoPs, and turbulence condition.



**Figure B2.** Experiment 2: Mean JoAs (circles) and model predictions (regression line, and shaded 95% prediction intervals) as a function of JoPs, and turbulence condition.



**Figure B3.** Experiment 3: Mean JoAs (circles) and model predictions (regression line, and shaded 95% prediction intervals) as a function of JoPs, and turbulence condition.

### Polynomial Analysis

To further assess whether non-linearities could be present in the relation between JoAs and JoPs, we used model comparison to test whether quadratic or cubic polynomial functions might better characterize the relation between JoAs and JoPs than a linear function.

For each experiment, we compared 3 nested hierarchical linear regressions. We obtained orthogonal linear, quadratic and cubic JoPs polynomials. Model 1 predicted JoAs as a function of the turbulence factor (coded as 1 = Turbulence, -1 = No Turbulence), and a linear effect of JoPs. Participants were modelled as random intercepts, and random slopes for the turbulence factor. Model 2 additionally included a quadratic effect of JoPs. Model 3 additionally included a cubic effect of JoPs. Full results tables are displayed below. Satterthwaite approximation for degrees of freedom (Kuznetsova et al., 2015) was used to obtain p-values, and 95% confidence intervals were obtained using the Wald method (Bates et al., 2014).

Model comparison using AIC showed that the linear model was preferred over the quadratic and cubic JoP models in Experiment 1 (AIC difference in favour of linear model over quadratic = 2.9, and over cubic = 6.7) and Experiment 3 (AIC differences in favour of linear over quadratic = 4, and over cubic = 6.8). For Experiment 2, the quadratic model was slightly favoured by AIC over linear (AIC difference = 1.6), though this is likely a spurious result. The linear model was still favoured over cubic (AIC difference = 2.5) in Experiment 2. Overall, the modelling results combined with observation of the data support the assumption of a linear relation between JoPs and JoAs.

## B.1. Experiment 1

**Table B1.** Experiment 1, Model 1 – Linear effect of JoPs.

	Estimate	SE	<i>t</i>	df	<i>p</i>	95 % CI	
						Lower	Upper
(Intercept)	0.52	0.01	37.00	21.97	< 0.001	0.49	0.55
Turbulence	-0.15	0.02	-9.58	28.87	< 0.001	-0.18	-0.12
JoP-Linear	3.82	0.22	17.18	443.66	< 0.001	3.39	4.26

**Table B2.** Experiment 1, Model 2 – Quadratic effect of JoPs.

	Estimate	SE	<i>t</i>	df	<i>p</i>	95 % CI	
						Lower	Upper
(Intercept)	0.52	0.01	36.55	21.79	< 0.001	0.49	0.55
Turbulence	-0.15	0.02	-9.37	28.68	< 0.001	-0.18	-0.12
JoP-Linear	3.86	0.23	16.85	438.59	< 0.001	3.41	4.31
JoP-Quadratic	0.15	0.17	0.89	503.80	0.374	-0.19	0.50

**Table B3.** Experiment 1, Model 3 – Cubic effect of JoPs.

	Estimate	SE	<i>t</i>	df	<i>p</i>	95 % CI	
						Lower	Upper
(Intercept)	0.52	0.01	36.57	21.76	< 0.001	0.49	0.55
Turbulence	-0.15	0.02	-9.38	28.83	< 0.001	-0.18	-0.12
JoP-Linear	3.86	0.23	16.84	436.96	< 0.001	3.41	4.31
JoP-Quadratic	0.16	0.18	0.92	489.68	0.356	-0.18	0.51
JoP-Cubic	0.05	0.16	0.31	546.42	0.757	-0.26	0.36

**Table B4.** Experiment 1: Model comparison.

	df	AIC
Model 1	7	-604.7
Model 2	8	-601.8
Model 3	9	-598.0

## B.2. Experiment 2

**Table B5.** Experiment 2, Model 1 – Linear effect of JoPs.

	Estimate	SE	<i>t</i>	df	<i>p</i>	95 % CI	
						Lower	Upper
(Intercept)	0.53	0.02	29.92	22.30	< 0.001	0.50	0.57
Turbulence	-0.16	0.02	-10.31	24.38	< 0.001	-0.19	-0.13

JoP-Linear      1.72      0.19      9.27      547.01      < 0.001      1.35      2.08

**Table B6.** Experiment 2, Model 2 – Quadratic effect of JoPs.

	Estimate	SE	<i>t</i>	df	<i>p</i>	95 % CI	
						Lower	Upper
(Intercept)	0.53	0.02	31.08	22.01	< 0.001	0.50	0.57
Turbulence	-0.16	0.02	-10.45	24.37	< 0.001	-0.19	-0.13
JoP-Linear	1.75	0.18	9.48	543.55	< 0.001	1.39	2.11
JoP-Quadratic	0.36	0.15	2.38	559.09	0.018	0.06	0.65

**Table B7.** Experiment 2, Model 3 – Cubic effect of JoPs.

	Estimate	SE	<i>t</i>	df	<i>p</i>	95 % CI	
						Lower	Upper
(Intercept)	0.53	0.02	31.06	22.00	< 0.001	0.50	0.57
Turbulence	-0.16	0.02	-10.45	24.37	< 0.001	-0.19	-0.13
JoP-Linear	1.75	0.19	9.45	541.63	< 0.001	1.39	2.11
JoP-Quadratic	0.36	0.15	2.38	558.11	0.018	0.06	0.65
JoP-Cubic	0.01	0.14	0.10	543.52	0.921	-0.26	0.29

**Table B8.** Experiment 2: Model comparison.

	df	AIC
Model 1	7	-578.0
Model 2	8	-579.6
Model 3	9	-575.5

### B.3. Experiment 3

**Table B9.** Experiment 3, Model 1 – Linear effect of JoPs.

	Estimate	SE	<i>t</i>	df	<i>p</i>	95 % CI	
						Lower	Upper
(Intercept)	0.59	0.01	46.96	20.81	< 0.001	0.56	0.61
Turbulence	-0.10	0.02	-5.74	21.66	< 0.001	-0.14	-0.07
JoP-Linear	4.05	0.17	23.50	670.19	< 0.001	3.71	4.38

**Table B10.** Experiment 3, Model 2 – Quadratic effect of JoPs.

	Estimate	SE	<i>t</i>	df	<i>p</i>	95 % CI	
						Lower	Upper
(Intercept)	0.59	0.01	46.91	20.80	< 0.001	0.56	0.61
Turbulence	-0.10	0.02	-5.74	21.66	< 0.001	-0.14	-0.07
JoP-Linear	4.05	0.17	23.46	672.86	< 0.001	3.71	4.38
JoP-Quadratic	0.01	0.15	0.05	761.74	0.958	-0.28	0.30

**Table B11.** Experiment 3, Model 3 – Cubic effect of JoPs.

	Estimate	SE	<i>t</i>	df	<i>p</i>	95 % CI	
						Lower	Upper
(Intercept)	0.59	0.01	47.03	20.80	< 0.001	0.56	0.61
Turbulence	-0.10	0.02	-5.73	21.66	< 0.001	-0.14	-0.07
JoP-Linear	4.04	0.17	23.45	671.96	< 0.001	3.70	4.38
JoP-Quadratic	0.01	0.15	0.05	760.75	0.964	-0.28	0.30
JoP-Cubic	-0.15	0.13	-1.16	771.93	0.248	-0.42	0.11

**Table B12.** Experiment 3: Model comparison.

	df	AIC
Model 1	7	-954.1
Model 2	8	-950.1
Model 3	9	-947.3



## Appendix C. Supplementary Data for Chapter 4

### C.1. Target N2

**Table C1.** Target N2 model: parameter estimates, with bootstrapped 95% confidence intervals. \* Based on Satterthwaite degrees of freedom estimation.

	Estimate	SE	<i>t</i>	df *	<i>p</i> *	95% CI	
						Lower	Upper
Intercept	6.77	0.86	7.88	23	< .001	5.15	8.27
Choice	-0.020	0.170	-0.11	23	0.91	-0.37	0.34
Priming	0.34	0.14	2.42	34	0.021	0.080	0.60
RTs (Z)	-1.98	0.12	-17.04	8325	< .001	-2.18	-1.77
Choice x Priming	-0.11	0.13	-0.84	37	0.41	-0.40	0.16
Choice x RTs	-0.071	0.12	-0.61	8494	0.54	-0.29	0.16
Priming x RTs	-0.25	0.12	-2.13	8449	0.033	-0.47	-0.0010
Choice x Priming x RTs	-0.099	0.12	-0.85	8353	0.39	-0.34	0.13

### C.2. Action CRN

#### C.2.1. Action CRN model

**Table C2.** Action CRN model: parameter estimates, with bootstrapped 95% confidence intervals. \* Based on Satterthwaite degrees of freedom estimation.

	Estimate	SE	<i>t</i>	df *	<i>p</i> *	95% CI	
						Lower	Upper
Intercept	3.23	0.53	6.05	23	< .001	2.27	4.21
Choice	0.13	0.15	0.86	21	0.40	-0.15	.41
Priming	-0.27	0.13	-2.12	24	0.044	-0.51	-0.020
RTs (Z)	-0.54	0.10	-5.55	8485	< .001	-0.72	-0.35
Choice x Priming	0.12	0.10	1.16	58	0.25	-0.090	0.32
Choice x RTs	0.17	0.10	1.73	8696	0.084	-0.023	0.36
Priming x RTs	0.37	0.10	3.85	8667	< .001	0.19	0.55
Choice x Priming x RTs	0.048	0.10	0.50	8458	0.62	-0.14	0.24

### C.2.2. Action CRN model with Target N2

**Table C3.** Action CRN model with Target N2: parameter estimates, with bootstrapped 95% confidence intervals. \* Based on Satterthwaite degrees of freedom estimation.

	Estimate	SE	<i>t</i>	df *	<i>p</i> *	95% CI	
						Lower	Upper
Intercept	3.23	0.53	6.05	23	< .001	2.15	4.30
Choice	0.13	0.150	0.88	21	0.39	-0.14	0.42
Priming	-0.24	0.13	-1.91	24	0.069	-0.48	0.011
RTs (Z)	-0.71	0.10	-7.27	8495	< .001	-0.91	-0.51
Target N2 (Z)	-0.94	0.09	-10.02	8775	< .001	-1.13	-0.75
Choice x Priming	0.11	0.10	1.08	55	0.28	-0.093	0.31
Choice x RTs	0.16	0.10	1.68	8693	0.093	-0.029	0.36
Priming x RTs	0.35	0.10	3.62	8663	< .001	0.16	0.56
Choice x Priming x RTs	0.040	0.10	0.42	8451	0.68	-0.13	0.21

### C.3. Outcome FRN

#### C.3.1. Outcome FRN model

**Table C4.** Outcome FRN model: parameter estimates, with bootstrapped 95% confidence intervals. \* Based on Satterthwaite degrees of freedom estimation.

	Estimate	SE	<i>t</i>	df *	<i>p</i> *	95 % CI	
						Lower	Upper
Intercept	-0.22	0.97	-0.23	23	0.82	-2.42	1.54
Choice	-0.015	0.17	-0.085	22.90	0.93	-0.36	0.31
Priming	-0.094	0.13	-0.75	42.30	0.46	-0.33	0.15
Choice x Priming	0.13	0.13	1.00	29	0.33	-0.16	0.39

**C.3.2. Outcome FRN model with Action CRN****Table C5.** Outcome FRN model with Action CRN: parameter estimates, with bootstrapped 95% confidence intervals. \* Based on Satterthwaite degrees of freedom estimation.

	Estimate	SE	<i>t</i>	<i>df</i> *	<i>p</i> *	95 % CI	
						Lower	Upper
Intercept	-0.22	0.97	-0.23	23	0.82	-2.20	1.57
Choice	-0.015	0.17	-0.085	23	0.93	-0.35	0.34
Priming	-0.095	0.13	-0.75	42	0.46	-0.33	0.17
Action CRN (Z)	-0.087	0.11	-0.76	8778	0.45	-0.31	0.12
Choice x Priming	0.13	0.13	1.00	29	0.32	-0.15	0.39

**C.4. Outcome N1****Table C6.** Outcome N1 model: parameter estimates, with bootstrapped 95% confidence intervals. \* Based on Satterthwaite degrees of freedom estimation.

	Estimate	SE	<i>t</i>	<i>df</i> *	<i>p</i> *	95 % CI	
						Lower	Upper
Intercept	-1.91	0.41	-4.63	22.90	< .001	-2.77	-1.14
Choice	-0.29	0.12	-2.38	25.30	0.025	-0.54	0.034
Priming	0.061	0.094	0.65	269	0.52	-0.14	0.25
Choice x Priming	-0.0042	0.11	-0.039	25.10	0.97	-0.22	0.20

## Appendix D. Supplementary Analyses for Chapter 5

### *Independent models.*

In addition to fitting a three-level multilevel regression model to all experiments simultaneously, we also fit separate two-level regression models to each experiment's data. We used the same coding for the variables for these models as the one presented in Chapter 5, but included all trials for each experiment. The by-subject varying effects (all average effects were allowed to vary within subjects) were modelled as uncorrelated because specifying correlated random effects led to non-convergence.

Tables D1-D7 present the results of these models, for each experiment. Effects of primary interest are highlighted in bold.

### **D.1. Group 1**

The results from the independent models fitted to experiments from Group 1 led to identical conclusions as the model presented in Chapter 5. Importantly, Congruency and Trial number exerted a positive influence on agency ratings across experiments, but action-outcome interval (AOI) did not. Congruency and Trial number interacted in all experiments, except in Experiment 2.

The lack of a Congruency by Trial interaction in Experiment 2 (**Table D2**) may result from two important differences in experimental design relative to other Group 1 experiments: there were 3 congruency conditions – congruent, neutral, and incongruent; and 3 AOIs. Notably, even though there were 3 AOIs, as in the experiments of Group 2, this factor did not influence agency ratings. While the neutral condition was excluded from the present analyses, it may have impacted on the learning of action-outcome contingencies, as well as on potential interactions between congruency and outcome identity. The significant effect of trial number shows that outcome identity was still driving agency ratings, but the slope of the increase in agency ratings across trials was reduced relative to other Group 1

experiments (see also **Figure 5.3**, p. 155). Therefore, this reduced the potential for dynamic changes in the effects of selection fluency on the sense of agency.

**Table D1.** Multilevel model of agency ratings for Experiment 1.

Average Effect	Estimate	SE	<i>t</i>	<i>p</i>	95% CI	
					Lower	Upper
Intercept	5.13	0.17	30.09	< 0.001	4.80	5.47
<b>Trial</b>	<b>5.69</b>	<b>0.89</b>	<b>6.43</b>	<b>&lt; 0.001</b>	<b>3.96</b>	<b>7.43</b>
Trial (quadratic)	-1.78	0.40	-4.44	< 0.001	-2.57	-1.00
<b>Congruency</b>	<b>0.23</b>	<b>0.09</b>	<b>2.48</b>	<b>0.019</b>	<b>0.05</b>	<b>0.41</b>
<b>AOI</b>	<b>0.00</b>	<b>0.02</b>	<b>0.17</b>	<b>0.861</b>	<b>-0.04</b>	<b>0.04</b>
<b>Trial x Congruency</b>	<b>0.94</b>	<b>0.44</b>	<b>2.13</b>	<b>0.044</b>	<b>0.07</b>	<b>1.81</b>
Congruency x AOI	0.04	0.04	1.01	0.321	-0.04	0.12

**Table D2.** Multilevel model of agency ratings for Experiment 2.

Average Effect	Estimate	SE	<i>t</i>	<i>p</i>	95% CI	
					Lower	Upper
Intercept	5.16	0.50	10.28	< 0.001	4.18	6.15
<b>Trial</b>	<b>3.99</b>	<b>1.34</b>	<b>2.98</b>	<b>0.013</b>	<b>1.36</b>	<b>6.61</b>
Trial (quadratic)	-1.29	0.60	-2.16	0.054	-2.46	-0.12
<b>Congruency</b>	<b>0.30</b>	<b>0.13</b>	<b>2.35</b>	<b>0.034</b>	<b>0.05</b>	<b>0.56</b>
<b>AOI</b>	<b>0.00</b>	<b>0.02</b>	<b>-0.13</b>	<b>0.895</b>	<b>-0.04</b>	<b>0.04</b>
<b>Trial x Congruency</b>	<b>0.08</b>	<b>0.75</b>	<b>0.10</b>	<b>0.919</b>	<b>-1.39</b>	<b>1.54</b>
Congruency x AOI	0.03	0.04	0.69	0.490	-0.05	0.11

**Table D3.** Multilevel model of agency ratings for Experiment 3.

Average Effect	Estimate	SE	<i>t</i>	<i>p</i>	95% CI	
					Lower	Upper
Intercept	5.98	0.23	25.46	< 0.001	5.52	6.44
<b>Trial</b>	<b>5.23</b>	<b>0.83</b>	<b>6.33</b>	<b>&lt; 0.001</b>	<b>3.61</b>	<b>6.85</b>
Trial (quadratic)	-2.25	0.36	-6.20	< 0.001	-2.96	-1.54
<b>Congruency</b>	<b>0.42</b>	<b>0.13</b>	<b>3.21</b>	<b>0.003</b>	<b>0.16</b>	<b>0.68</b>
<b>AOI</b>	<b>0.03</b>	<b>0.02</b>	<b>1.88</b>	<b>0.061</b>	<b>0.00</b>	<b>0.07</b>
<b>Trial x Congruency</b>	<b>0.82</b>	<b>0.42</b>	<b>1.96</b>	<b>0.063</b>	<b>0.00</b>	<b>1.64</b>
Congruency x AOI	-0.04	0.04	-1.12	0.263	-0.11	0.03

**Table D4.** Multilevel model of agency ratings for Experiment 4.

Average Effect	Estimate	SE	<i>t</i>	<i>p</i>	95% CI	
					Lower	Upper
Intercept	6.44	0.30	21.55	< 0.001	5.85	7.03
<b>Trial</b>	<b>7.47</b>	<b>1.11</b>	<b>6.71</b>	<b>&lt; 0.001</b>	<b>5.29</b>	<b>9.65</b>
Trial (quadratic)	-3.58	0.55	-6.48	< 0.001	-4.66	-2.50
<b>Congruency</b>	<b>0.34</b>	<b>0.13</b>	<b>2.64</b>	<b>0.014</b>	<b>0.09</b>	<b>0.58</b>
<b>AOI</b>	<b>0.00</b>	<b>0.01</b>	<b>0.73</b>	<b>0.465</b>	<b>-0.01</b>	<b>0.02</b>
<b>Trial x Congruency</b>	<b>0.74</b>	<b>0.31</b>	<b>2.39</b>	<b>0.025</b>	<b>0.13</b>	<b>1.34</b>
Congruency x AOI	0.03	0.01	2.22	0.027	0.00	0.05

## D.2. Group 2

The independent model fits presented here further corroborate the results obtained from the three-level multilevel model presented in the main article. Congruency and AOI both exerted noticeable effects on SoA. Importantly, Trial number and Trial by Congruency effects were not reliable in the studies in Group 2. There was a small effect of Trial number in Experiment 6 (**Table D6**; see also **Figure 5.3**, p. 155). The slight increase in agency ratings over trials found for this experiment may be due to the fact that it had the longest block length of these experiments (**Table 5.1**, p. 154). This suggests that focusing on AOI may slow down the learning of action-outcome contingencies, thus requiring longer before outcome identity influences SoA.

**Table D5.** Multilevel model of agency ratings for Experiment 5.

Average Effect	Estimate	SE	<i>t</i>	<i>p</i>	95% CI	
					Lower	Upper
Intercept	4.31	0.16	26.58	< 0.001	3.99	4.62
<b>Trial</b>	<b>0.47</b>	<b>0.32</b>	<b>1.46</b>	<b>0.158</b>	<b>-0.16</b>	<b>1.11</b>
Trial (quadratic)	-0.43	0.32	-1.33	0.197	-1.05	0.20
<b>Congruency</b>	<b>0.20</b>	<b>0.07</b>	<b>2.91</b>	<b>0.006</b>	<b>0.06</b>	<b>0.33</b>
<b>AOI</b>	<b>-0.21</b>	<b>0.08</b>	<b>-2.65</b>	<b>0.015</b>	<b>-0.37</b>	<b>-0.06</b>
<b>Trial x Congruency</b>	<b>-0.90</b>	<b>0.49</b>	<b>-1.83</b>	<b>0.082</b>	<b>-1.86</b>	<b>0.07</b>
Congruency x AOI	0.02	0.03	0.68	0.499	-0.04	0.09

**Table D6.** Multilevel model of agency ratings for Experiment 6.

Average Effect	Estimate	SE	<i>t</i>	<i>p</i>	95% CI	
					Lower	Upper
Intercept	4.93	0.13	38.32	< 0.001	4.68	5.18
<b>Trial</b>	<b>0.99</b>	<b>0.39</b>	<b>2.50</b>	<b>0.027</b>	<b>0.21</b>	<b>1.76</b>
Trial (quadratic)	0.12	0.24	0.52	0.601	-0.34	0.59
<b>Congruency</b>	<b>0.31</b>	<b>0.11</b>	<b>2.75</b>	<b>0.016</b>	<b>0.09</b>	<b>0.53</b>
<b>AOI</b>	<b>-0.33</b>	<b>0.08</b>	<b>-4.07</b>	<b>0.001</b>	<b>-0.49</b>	<b>-0.17</b>
<b>Trial x Congruency</b>	<b>-1.03</b>	<b>0.54</b>	<b>-1.90</b>	<b>0.080</b>	<b>-2.09</b>	<b>0.03</b>
Congruency x AOI	0.01	0.03	0.28	0.785	-0.05	0.07

**Table D7.** Multilevel model of agency ratings for Experiment 7.

Average Effect	Estimate	SE	<i>t</i>	<i>p</i>	95% CI	
					Lower	Upper
Intercept	4.31	0.35	12.21	< 0.001	3.62	5.00
<b>Trial</b>	<b>0.41</b>	<b>0.52</b>	<b>0.78</b>	<b>0.452</b>	<b>-0.62</b>	<b>1.44</b>
Trial (quadratic)	-0.32	0.35	-0.91	0.364	-1.00	0.36
<b>Congruency</b>	<b>0.33</b>	<b>0.13</b>	<b>2.43</b>	<b>0.015</b>	<b>0.06</b>	<b>0.59</b>
<b>AOI</b>	<b>-0.30</b>	<b>0.14</b>	<b>-2.08</b>	<b>0.062</b>	<b>-0.59</b>	<b>-0.02</b>
<b>Trial x Congruency</b>	<b>0.72</b>	<b>0.81</b>	<b>0.89</b>	<b>0.390</b>	<b>-0.86</b>	<b>2.30</b>
Congruency x AOI	0.08	0.07	1.14	0.255	-0.06	0.22

## References

- Aarts, H., Custers, R., & Wegner, D. M. (2005). On the inference of personal authorship: enhancing experienced agency by priming effect information. *Consciousness and Cognition*, 14(3), 439–458.  
<http://doi.org/10.1016/j.concog.2004.11.001>
- Aarts, K., De Houwer, J., & Pourtois, G. (2013). Erroneous and correct actions have a different affective valence: Evidence from ERPs. *Emotion*, 13(5), 960–973.  
<http://doi.org/10.1037/a0032808>
- Alter, A. L., & Oppenheimer, D. M. (2009). Uniting the Tribes of Fluency to Form a Metacognitive Nation. *Personality and Social Psychology Review*, 13(3), 219–235. <http://doi.org/10.1177/1088868309341564>
- Amodio, D. M., & Frith, C. D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nature Reviews. Neuroscience*, 7(4), 268–277.  
<http://doi.org/10.1038/nrn1884>
- Arango-Muñoz, S. (2010). Two Levels of Metacognition. *Philosophia*, 39, 71–82.  
<http://doi.org/10.1007/s11406-010-9279-0>
- Atas, A., & Cleeremans, A. (2015). The temporal dynamic of automatic inhibition of irrelevant actions. *Journal of Experimental Psychology: Human Perception and Performance*, 41(2), 289–305. <http://doi.org/10.1037/a0038654>
- Atas, A., Desender, K., Gevers, W., & Cleeremans, A. (2015). Dissociating Perception From Action During Conscious and Unconscious Conflict Adaptation. *Journal of Experimental Psychology. Learning, Memory, and Cognition*. <http://doi.org/10.1037/xlm0000206>
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(4), 390–412. <http://doi.org/10.1016/j.jml.2007.12.005>
- Baess, P., Widmann, A., Roze, A., Schröger, E., & Jacobsen, T. (2009). Attenuated human auditory middle latency response and evoked 40-Hz response to self-initiated sounds. *The European Journal of Neuroscience*, 29(7), 1514–1521.  
<http://doi.org/10.1111/j.1460-9568.2009.06683.x>



- Bagiella, E., Sloan, R. P., & Heitjan, D. F. (2000). Mixed-effects models in psychophysiology. *Psychophysiology*, 37(1), 13–20.  
<http://doi.org/10.1111/1469-8986.3710013>
- Bandura, A. (1991). Social cognitive theory of self-regulation. *Organizational Behavior and Human Decision Processes*, 50(2), 248–287.  
[http://doi.org/10.1016/0749-5978\(91\)90022-L](http://doi.org/10.1016/0749-5978(91)90022-L)
- Bandura, A. (1999). Moral disengagement in the perpetration of inhumanities. *Personality and Social Psychology Review*, 3(3), 193–209.
- Bandura, A., Underwood, B., & Fromson, M. E. (1975). Disinhibition of aggression through diffusion of responsibility and dehumanization of victims. *Journal of Research in Personality*, 9(4), 253–269. [http://doi.org/10.1016/0092-6566\(75\)90001-X](http://doi.org/10.1016/0092-6566(75)90001-X)
- Barlas, Z., & Obhi, S. S. (2013). Freedom, choice, and the sense of agency. *Frontiers in Human Neuroscience*, 7, 514.  
<http://doi.org/10.3389/fnhum.2013.00514>
- Baron, J., & Ritov, I. (2004). Omission bias, individual differences, and normality. *Organizational Behavior and Human Decision Processes*, 94(2), 74–85.  
<http://doi.org/10.1016/j.obhdp.2004.03.003>
- Bartholow, B. D., Pearson, M. A., Dickter, C. L., Sher, K. J., Fabiani, M., & Gratton, G. (2005). Strategic control and medial frontal negativity: Beyond errors and response conflict. *Psychophysiology*, 42(1), 33–42.  
<http://doi.org/10.1111/j.1469-8986.2005.00258.x>
- Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., & Dai, B. (2014). lme4: Linear mixed-effects models using Eigen and S4 (Version 1.1-7). Retrieved from  
<http://cran.r-project.org/web/packages/lme4/index.html>
- Bednark, J. G., & Franz, E. A. (2014). Agency attribution: event-related potentials and outcome monitoring. *Experimental Brain Research*, 1–10.  
<http://doi.org/10.1007/s00221-014-3821-4>
- Belayachi, S., & Van der Linden, M. (2010). Feeling of doing in obsessive–compulsive checking. *Consciousness and Cognition*, 19(2), 534–546.  
<http://doi.org/10.1016/j.concog.2010.02.001>

- Berberian, B., Sarrazin, J.-C., Le Blaye, P., & Haggard, P. (2012). Automation technology and sense of control: a window on human agency. *PloS One*, 7(3), e34075. <http://doi.org/10.1371/journal.pone.0034075>
- Beste, C., Saft, C., Andrich, J., Gold, R., & Falkenstein, M. (2008). Stimulus-Response Compatibility in Huntington's Disease: A Cognitive-Neurophysiological Analysis. *Journal of Neurophysiology*, 99(3), 1213–1223. <http://doi.org/10.1152/jn.01152.2007>
- Blakemore, S. J., Frith, C. D., & Wolpert, D. M. (1999). Spatio-temporal prediction modulates the perception of self-produced stimuli. *Journal of Cognitive Neuroscience*, 11(5), 551–559.
- Blakemore, S. J., Wolpert, D. M., & Frith, C. D. (1998). Central cancellation of self-produced tickle sensation. *Nature Neuroscience*, 1(7), 635–640. <http://doi.org/10.1038/2870>
- Blakemore, S. J., Wolpert, D. M., & Frith, C. D. (2002). Abnormalities in the awareness of action. *Trends in Cognitive Sciences*, 6(6), 237–242. [http://doi.org/10.1016/S1364-6613\(02\)01907-1](http://doi.org/10.1016/S1364-6613(02)01907-1)
- Boldt, A., & Yeung, N. (2015). Shared Neural Markers of Decision Confidence and Error Detection. *The Journal of Neuroscience*, 35(8), 3478–3484. <http://doi.org/10.1523/JNEUROSCI.0797-14.2015>
- Botvinick, M. M. (2007). Conflict monitoring and decision making: Reconciling two perspectives on anterior cingulate function. *Cognitive, Affective, & Behavioral Neuroscience*, 7(4), 356–366. <http://doi.org/10.3758/CABN.7.4.356>
- Botvinick, M. M., & Braver, T. (2015). Motivation and Cognitive Control: From Behavior to Neural Mechanism. *Annual Review of Psychology*, 66(1), 83–113. <http://doi.org/10.1146/annurev-psych-010814-015044>
- Bradley, G. L. (1995). Group influences upon preferences for personal protection: A simulation study. *Journal of Safety Research*, 26(2), 99–105. [http://doi.org/10.1016/0022-4375\(95\)00009-F](http://doi.org/10.1016/0022-4375(95)00009-F)
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10(4), 433–436.
- Braver, T. S. (2012). The variable nature of cognitive control: a dual mechanisms framework. *Trends in Cognitive Sciences*, 16(2), 106–113. <http://doi.org/10.1016/j.tics.2011.12.010>

- Carter, C. S., Braver, T. S., Barch, D. M., Botvinick, M. M., Noll, D., & Cohen, J. D. (1998). Anterior Cingulate Cortex, Error Detection, and the Online Monitoring of Performance. *Science*, *280*(5364), 747–749.
- Caspar, E. A., Christensen, J. F., Cleeremans, A., & Haggard, P. (2016). Coercion Changes the Sense of Agency in the Human Brain. *Current Biology*, *26*(5), 585–592. <http://doi.org/10.1016/j.cub.2015.12.067>
- Caspar, E. A., Cleeremans, A., & Haggard, P. (2015). The relationship between human agency and embodiment. *Consciousness and Cognition*, *33*, 226–236. <http://doi.org/10.1016/j.concog.2015.01.007>
- Cavazzana, A., Penolazzi, B., Begliomini, C., & Bisiacchi, P. S. (2015). Neural underpinnings of the “agent brain”: new evidence from transcranial direct current stimulation. *European Journal of Neuroscience*, *42* (3), 1889–1894. <http://doi.org/10.1111/ejn.12937>
- Chambon, V., Filevich, E., & Haggard, P. (2014). What is the Human Sense of Agency, and is it Metacognitive? In S. M. Fleming & C. D. Frith (Eds.), *The Cognitive Neuroscience of Metacognition* (pp. 321–342). Springer Berlin Heidelberg. Retrieved from [http://link.springer.com/chapter/10.1007/978-3-642-45190-4\\_14](http://link.springer.com/chapter/10.1007/978-3-642-45190-4_14)
- Chambon, V., & Haggard, P. (2012). Sense of control depends on fluency of action selection, not motor performance. *Cognition*. <http://doi.org/10.1016/j.cognition.2012.07.011>
- Chambon, V., Moore, J. W., & Haggard, P. (2014). TMS stimulation over the inferior parietal cortex disrupts prospective sense of agency. *Brain Structure and Function*, 1–13. <http://doi.org/10.1007/s00429-014-0878-6>
- Chambon, V., Sidarus, N., & Haggard, P. (2014). From action intentions to action effects: how does the sense of agency come about? *Frontiers in Human Neuroscience*, *8*, 320. <http://doi.org/10.3389/fnhum.2014.00320>
- Chambon, V., Wenke, D., Fleming, S. M., Prinz, W., & Haggard, P. (2013). An Online Neural Substrate for Sense of Agency. *Cerebral Cortex*, *23*(5), 1031–1037. <http://doi.org/10.1093/cercor/bhs059>
- Charles, L., Van Opstal, F., Marti, S., & Dehaene, S. (2013). Distinct brain mechanisms for conscious versus subliminal error detection. *NeuroImage*, *73*,

- 80–94. <http://doi.org/10.1016/j.neuroimage.2013.01.054>
- Chekroun, P., & Brauer, M. (2002). The bystander effect and social control behavior: the effect of the presence of others on people's reactions to norm violations. *European Journal of Social Psychology*, 32(6), 853–867.  
<http://doi.org/10.1002/ejsp.126>
- Cisek, P. (2007). Cortical mechanisms of action selection: the affordance competition hypothesis. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 362(1485), 1585–1599.  
<http://doi.org/10.1098/rstb.2007.2054>
- Cohen, M. X., & Donner, T. H. (2013). Midfrontal conflict-related theta-band power reflects neural oscillations that predict behavior. *Journal of Neurophysiology*, 110(12), 2752–2763. <http://doi.org/10.1152/jn.00479.2013>
- Csikszentmihalyi, M. (1990). *Flow: The Psychology of Optimal Experience*. Harper & Row.
- Damen, T. G. E., van Baaren, R. B., & Dijksterhuis, A. (2014). You should read this! Perceiving and acting upon action primes influences one's sense of agency. *Journal of Experimental Social Psychology*, 50, 21–26.  
<http://doi.org/10.1016/j.jesp.2013.09.003>
- Darley, J. M., & Latane, B. (1968). Bystander Intervention in Emergencies: Diffusion of Responsibility. *Journal of Personality*, 377–383.
- David, N., Newen, A., & Vogeley, K. (2008). The “sense of agency” and its underlying cognitive and neural mechanisms. *Consciousness and Cognition*, 17(2), 523–534. <http://doi.org/10.1016/j.concog.2008.03.004>
- Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1), 9–21.  
<http://doi.org/10.1016/j.jneumeth.2003.10.009>
- Demanet, J., De Baene, W., Arrington, C. M., & Brass, M. (2013). Biasing free choices: The role of the rostral cingulate zone in intentional control. *NeuroImage*. <http://doi.org/10.1016/j.neuroimage.2013.01.052>
- Demanet, J., Muhle-Karbe, P. S., Lynn, M. T., Blotenberg, I., & Brass, M. (2013). Power to the will: How exerting physical effort boosts the sense of agency.

- Cognition*, 129(3), 574–578. <http://doi.org/10.1016/j.cognition.2013.08.020>
- Desantis, A., Weiss, C., Schutz-Bosbach, S., & Waszak, F. (2012). Believing and Perceiving: Authorship Belief Modulates Sensory Attenuation. *PLoS ONE*, 7(5). <http://doi.org/10.1371/journal.pone.0037959>
- Desender, K., Opstal, F. V., & Bussche, E. V. den. (2014). Feeling the Conflict The Crucial Role of Conflict Experience in Adaptation. *Psychological Science*, 25(3), 675–683. <http://doi.org/10.1177/0956797613511468>
- Desender, K., Van Lierde, E., & Van den Bussche, E. (2013). Comparing Conscious and Unconscious Conflict Adaptation. *PLoS ONE*, 8(2). <http://doi.org/10.1371/journal.pone.0055976>
- Dewey, J. A., & Knoblich, G. (2014). Do Implicit and Explicit Measures of the Sense of Agency Measure the Same Thing? *PLoS ONE*, 9(10), e110118. <http://doi.org/10.1371/journal.pone.0110118>
- Dewey, J. A., Seiffert, A. E., & Carr, T. H. (2010). Taking credit for success: The phenomenology of control in a goal-directed task. *Consciousness and Cognition*, 19(1), 48–62. <http://doi.org/10.1016/j.concog.2009.09.007>
- de Wit, S., & Dickinson, A. (2009). Associative theories of goal-directed behaviour: a case for animal-human translational models. *Psychological Research*, 73(4), 463–476. <http://doi.org/10.1007/s00426-009-0230-6>
- Donkers, F. C. L., & van Boxtel, G. J. M. (2004). The N2 in go/no-go tasks reflects conflict monitoring not response inhibition. *Brain and Cognition*, 56(2), 165–176. <http://doi.org/10.1016/j.bandc.2004.04.005>
- Dreisbach, G., & Fischer, R. (2011). If it's hard to read... try harder! Processing fluency as signal for effort adjustments. *Psychological Research*, 75(5), 376–383. <http://doi.org/10.1007/s00426-010-0319-y>
- Dreisbach, G., & Fischer, R. (2012). Conflicts as aversive signals. *Brain and Cognition*, 78(2), 94–98. <http://doi.org/10.1016/j.bandc.2011.12.003>
- Ebert, J. P., & Wegner, D. M. (2010). Time warp: authorship shapes the perceived timing of actions and events. *Consciousness and Cognition*, 19(1), 481–489. <http://doi.org/10.1016/j.concog.2009.10.002>
- Efron, B., & Tibshirani, R. J. (1994). *An Introduction to the Bootstrap*. CRC Press.
- Egner, T. (2007). Congruency sequence effects and cognitive control. *Cognitive*,

- Affective, & Behavioral Neuroscience*, 7(4), 380–390.  
<http://doi.org/10.3758/CABN.7.4.380>
- Eimer, M., & Schlaghecken, F. (1998). Effects of masked stimuli on motor activation: behavioral and electrophysiological evidence. *Journal of Experimental Psychology. Human Perception and Performance*, 24(6), 1737–1747.
- Eimer, M., & Schlaghecken, F. (2003). Response facilitation and inhibition in subliminal priming. *Biological Psychology*, 64(1-2), 7–26.
- Elsner, B., & Hommel, B. (2001). Effect anticipation and action control. *Journal of Experimental Psychology. Human Perception and Performance*, 27(1), 229–240.
- Elsner, B., Hommel, B., Mentschel, C., Drzezga, A., Prinz, W., Conrad, B., & Siebner, H. (2002). Linking Actions and Their Perceivable Consequences in the Human Brain. *NeuroImage*, 17(1), 364–372.  
<http://doi.org/10.1006/nimg.2002.1162>
- Endrass, T., Klawohn, J., Gruetzmänn, R., Ischebeck, M., & Kathmann, N. (2012). Response-related negativities following correct and incorrect responses: Evidence from a temporospatial principal component analysis. *Psychophysiology*, 49(6), 733–743.  
<http://doi.org/10.1111/j.1469-8986.2012.01365.x>
- Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, 16(1), 143–149. <http://doi.org/10.3758/BF03203267>
- Eriksen, C. W., & Schultz, D. W. (1979). Information processing in visual search: A continuous flow conception and experimental results. *Perception & Psychophysics*, 25(4), 249–263. <http://doi.org/10.3758/BF03198804>
- Evans, N., Gale, S., Schurger, A., & Blanke, O. (2015). Visual Feedback Dominates the Sense of Agency for Brain-Machine Actions. *PLoS ONE*, 10(6), e0130019. <http://doi.org/10.1371/journal.pone.0130019>
- Farrer, C., Bouchereau, M., Jeannerod, M., & Franck, N. (2008). Effect of distorted visual feedback on the sense of agency. *Behavioural Neurology*, 19(1), 53–57.
- Farrer, C., Franck, N., Georgieff, N., Frith, C. D., Decety, J., & Jeannerod, M. (2003). Modulating the experience of agency: a positron emission tomography study.

- NeuroImage*, 18(2), 324–333.
- Farrer, C., Frey, S. H., Van Horn, J. D., Tunik, E., Turk, D., Inati, S., & Grafton, S. T. (2008). The Angular Gyrus Computes Action Awareness Representations. *Cerebral Cortex*, 18(2), 254–261. <http://doi.org/10.1093/cercor/bhm050>
- Farrer, C., & Frith, C. D. (2002). Experiencing oneself vs another person as being the cause of an action: the neural correlates of the experience of agency. *NeuroImage*, 15(3), 596–603. <http://doi.org/10.1006/nimg.2001.1009>
- Farrer, C., Valentin, G., & Hupé, J. M. (2013). The time windows of the sense of agency. *Consciousness and Cognition*, 22(4), 1431–1441. <http://doi.org/10.1016/j.concog.2013.09.010>
- Fink, G. R., Marshall, J. C., Halligan, P. W., Frith, C. D., Driver, J., Frackowiak, R. S., & Dolan, R. (1999). The neural consequences of conflict between intention and the senses. *Brain: A Journal of Neurology*, 122 ( Pt 3), 497–512.
- Fleming, S. M., & Dolan, R. J. (2012). The neural basis of metacognitive ability. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1594), 1338–1349. <http://doi.org/10.1098/rstb.2011.0417>
- Fleming, S. M., & Dolan, R. J. (2014). The Neural Basis of Metacognitive Ability. In S. M. Fleming & C. D. Frith (Eds.), *The Cognitive Neuroscience of Metacognition* (pp. 245–265). Springer Berlin Heidelberg. Retrieved from [http://link.springer.com/chapter/10.1007/978-3-642-45190-4\\_11](http://link.springer.com/chapter/10.1007/978-3-642-45190-4_11)
- Fleming, S. M., Maniscalco, B., Ko, Y., Amendi, N., Ro, T., & Lau, H. (2014). Action-Specific Disruption of Perceptual Confidence. *Psychological Science*, 0956797614557697. <http://doi.org/10.1177/0956797614557697>
- Fleming, S. M., Mars, R. B., Gladwin, T. E., & Haggard, P. (2009). When the Brain Changes Its Mind: Flexibility of Action Selection in Instructed and Free Choices. *Cerebral Cortex*, 19(10), 2352–2360. <http://doi.org/10.1093/cercor/bhn252>
- Fleming, S. M., Weil, R. S., Nagy, Z., Dolan, R. J., & Rees, G. (2010). Relating introspective accuracy to individual differences in brain structure. *Science (New York, N.Y.)*, 329(5998), 1541–1543. <http://doi.org/10.1126/science.1191883>
- Flowers, J. H. (1990). Priming effects in perceptual classification. *Perception & Psychophysics*, 47(2), 135–148. <http://doi.org/10.3758/BF03205978>
- Forsyth, D. R., Zyzanski, L. E., & Giammanco, C. A. (2002). Responsibility

- diffusion in cooperative collectives. *Personality and Social Psychology Bulletin*, 28(1), 54–65.
- Frank, M. J., Wroch, B. S., & Curran, T. (2005). Error-Related Negativity Predicts Reinforcement Learning and Conflict Biases. *Neuron*, 47(4), 495–501. <http://doi.org/10.1016/j.neuron.2005.06.020>
- Frith, C. D. (2005). The self in action: Lessons from delusions of control. *Consciousness and Cognition*, 14(4), 752–770. <http://doi.org/10.1016/j.concog.2005.04.002>
- Frith, C. D. (2014). Action, agency and responsibility. *Neuropsychologia*, 55, 137–142. <http://doi.org/10.1016/j.neuropsychologia.2013.09.007>
- Fritz, J., & Dreisbach, G. (2013). Conflicts as aversive signals: Conflict priming increases negative judgments for neutral stimuli. *Cognitive, Affective, & Behavioral Neuroscience*, 13(2), 311–317. <http://doi.org/10.3758/s13415-012-0147-1>
- Gallagher, S. (2012). Multiple aspects in the sense of agency. *New Ideas in Psychology*, 30(1), 15–31. <http://doi.org/10.1016/j.newideapsych.2010.03.003>
- Gehring, W. J., & Willoughby, A. R. (2002). The Medial Frontal Cortex and the Rapid Processing of Monetary Gains and Losses. *Science*, 295(5563), 2279–2282. <http://doi.org/10.1126/science.1066893>
- Gelman, A., & Hill, J. (2006). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge: Cambridge University Press. Retrieved from <http://ebooks.cambridge.org/ref/id/CBO9780511790942>
- Gelman, A., & Su, Y.-S. (2015). *arm: Data Analysis Using Regression and Multilevel/Hierarchical Models*. Retrieved from <http://CRAN.R-project.org/package=arm>
- Gentsch, A., Kathmann, N., & Schütz-Bosbach, S. (2012). Reliability of sensory predictions determines the experience of self-agency. *Behavioural Brain Research*, 228(2), 415–422. <http://doi.org/10.1016/j.bbr.2011.12.029>
- Gentsch, A., & Schütz-Bosbach, S. (2011). I Did It: Unconscious Expectation of Sensory Consequences Modulates the Experience of Self-agency and Its Functional Signature. *Journal of Cognitive Neuroscience*, 23(12), 1–12. [http://doi.org/10.1162/jocn\\_a\\_00012](http://doi.org/10.1162/jocn_a_00012)



- Gentsch, A., Schütz-Bosbach, S., Endrass, T., & Kathmann, N. (2012). Dysfunctional forward model mechanisms and aberrant sense of agency in obsessive-compulsive disorder. *Biological Psychiatry*, 71(7), 652–659.  
<http://doi.org/10.1016/j.biopsych.2011.12.022>
- Gentsch, A., Ullsperger, P., & Ullsperger, M. (2009). Dissociable medial frontal negativities from a common monitoring system for self- and externally caused failure of goal achievement. *NeuroImage*, 47(4), 2023–2030.  
<http://doi.org/10.1016/j.neuroimage.2009.05.064>
- Gratton, G., Coles, M. G., & Donchin, E. (1992). Optimizing the use of information: strategic control of activation of responses. *Journal of Experimental Psychology. General*, 121(4), 480–506.
- Gray, R., Beilock, S. L., & Carr, T. H. (2007). “As soon as the bat met the ball, I knew it was gone”: Outcome prediction, hindsight bias, and the representation and control of action in expert and novice baseball players. *Psychonomic Bulletin & Review*, 14(4), 669–675. <http://doi.org/10.3758/BF03196819>
- Green, D. M., & Swets, J. A. (1966). *Signal Detection Theory and Psychophysics*. New York: Wiley.
- Grosjean, M., Zwickel, J., & Prinz, W. (2009). Acting while perceiving: assimilation precedes contrast. *Psychological Research*, 73(1), 3–13.  
<http://doi.org/10.1007/s00426-008-0146-6>
- Grützmann, R., Endrass, T., Klawohn, J., & Kathmann, N. (2014). Response accuracy rating modulates ERN and Pe amplitudes. *Biological Psychology*, 96, 1–7. <http://doi.org/10.1016/j.biopsycho.2013.10.007>
- Grützmann, R., Riesel, A., Klawohn, J., Kathmann, N., & Endrass, T. (2014). Complementary modulation of N2 and CRN by conflict frequency. *Psychophysiology*, 51(8), 761–772. <http://doi.org/10.1111/psyp.12222>
- Haggard, P., & Chambon, V. (2012). Sense of agency. *Current Biology*, 22(10), R390–392. <http://doi.org/10.1016/j.cub.2012.02.040>
- Haggard, P., Clark, S., & Kalogeras, J. (2002). Voluntary action and conscious awareness. *Nature Neuroscience*, 5(4), 382–385. <http://doi.org/10.1038/nn827>
- Haggard, P., & Tsakiris, M. (2009). The Experience of Agency. *Current Directions in Psychological Science*, 18(4), 242–246.

- <http://doi.org/10.1111/j.1467-8721.2009.01644.x>
- Hajcak, G., & Foti, D. (2008). Errors Are Aversive Defensive Motivation and the Error-Related Negativity. *Psychological Science*, 19(2), 103–108.  
<http://doi.org/10.1111/j.1467-9280.2008.02053.x>
- Hajcak, G., McDonald, N., & Simons, R. F. (2004). Error-related psychophysiology and negative affect. *Brain and Cognition*, 56(2), 189–197.  
<http://doi.org/10.1016/j.bandc.2003.11.001>
- Hakstian, A. R., Suedfeld, P., Ballard, E. J., & Rank, D. S. (1986). The Ascription of Responsibility Questionnaire: Development and Empirical Extensions. *Journal of Personality Assessment*, 50(2), 229–247.  
[http://doi.org/10.1207/s15327752jpa5002\\_10](http://doi.org/10.1207/s15327752jpa5002_10)
- Hauser, M., Moore, J. W., de Millas, W., Gallinat, J., Heinz, A., Haggard, P., & Voss, M. (2011). Sense of agency is altered in patients with a putative psychotic prodrome. *Schizophrenia Research*, 126(1-3), 20–27.  
<http://doi.org/10.1016/j.schres.2010.10.031>
- Heath, C., & Tversky, A. (1991). Preference and belief: Ambiguity and competence in choice under uncertainty. *Journal of Risk and Uncertainty*, 4(1), 5–28.  
<http://doi.org/10.1007/BF00057884>
- Holroyd, C. B., & Coles, M. G. H. (2002). The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, 109(4), 679–709.
- Holroyd, C. B., Krigolson, O. E., Baker, R., Lee, S., & Gibson, J. (2009). When is an error not a prediction error? An electrophysiological investigation. *Cognitive, Affective, & Behavioral Neuroscience*, 9(1), 59–70.  
<http://doi.org/10.3758/CABN.9.1.59>
- Holroyd, C. B., Pakzad-Vaezi, K. L., & Krigolson, O. E. (2008). The feedback correct-related positivity: Sensitivity of the event-related brain potential to unexpected positive feedback. *Psychophysiology*, 45(5), 688–697.  
<http://doi.org/10.1111/j.1469-8986.2008.00668.x>
- Holroyd, C. B., & Yeung, N. (2012). Motivation of extended behaviors by anterior cingulate cortex. *Trends in Cognitive Sciences*, 16(2), 122–128.  
<http://doi.org/10.1016/j.tics.2011.12.008>

- Horan, W. P., Foti, D., Hajcak, G., Wynn, J. K., & Green, M. F. (2012). Impaired Neural Response To Internal But Not External Feedback In Schizophrenia. *Psychological Medicine*, 42(8), 1637–1647.  
<http://doi.org/10.1017/S0033291711002819>
- Hughes, G., Desantis, A., & Waszak, F. (2012). Mechanisms of Intentional Binding and Sensory Attenuation: The Role of Temporal Prediction, Temporal Control, Identity Prediction, and Motor Prediction. *Psychological Bulletin*.  
<http://doi.org/10.1037/a0028566>
- Hughes, G., Desantis, A., & Waszak, F. (2013). Attenuation of auditory N1 results from identity-specific action-effect prediction. *The European Journal of Neuroscience*. <http://doi.org/10.1111/ejn.12120>
- Hughes, G., Velmans, M., & De Fockert, J. (2009). Unconscious priming of a no-go response. *Psychophysiology*, 46(6), 1258–1269. <http://doi.org/10.1111/j.1469-8986.2009.00873.x>
- Hughes, G., & Waszak, F. (2011). ERP correlates of action effect prediction and visual sensory attenuation in voluntary action. *NeuroImage*, 56(3), 1632–1640.  
<http://doi.org/10.1016/j.neuroimage.2011.02.057>
- Hume, D. (1740). *A treatise of human nature*. Oxford: Oxford University Press.
- Iyengar, S. S., & Lepper, M. R. (2000). When choice is demotivating: Can one desire too much of a good thing? *Journal of Personality and Social Psychology*, 79(6), 995–1006. <http://doi.org/10.1037/0022-3514.79.6.995>
- James, W. (1890). *The principles of psychology*. Macmillan.
- Javadi, A.-H. (2015). Modulation of the pre-SMA reduces sense of agency. *European Journal of Neuroscience*, n/a–n/a. <http://doi.org/10.1111/ejn.12954>
- Jiang, J., van Gaal, S., Bailey, K., Chen, A., & Zhang, Q. (2013). Electrophysiological correlates of block-wise strategic adaptations to consciously and unconsciously triggered conflict. *Neuropsychologia*, 51(13), 2791–2798.  
<http://doi.org/10.1016/j.neuropsychologia.2013.09.020>
- Jonides, J., & Mack, R. (1984). On the cost and benefit of cost and benefit. *Psychological Bulletin*, 96(1), 29–44. <http://doi.org/10.1037/0033-2909.96.1.29>
- Karau, S. J., & Williams, K. D. (1993). Social loafing: A meta-analytic review and theoretical integration. *Journal of Personality and Social Psychology*, 65(4),

- 681–706. <http://doi.org/10.1037/0022-3514.65.4.681>
- Khalighinejad, N., Di Costa, S., & Haggard, P. (2016). Endogenous Action Selection Processes in Dorsolateral Prefrontal Cortex Contribute to Sense of Agency: a Meta-Analysis of tDCS Studies of “Intentional Binding.” *Brain Stimulation*. <http://doi.org/10.1016/j.brs.2016.01.005>
- Khalighinejad, N., & Haggard, P. (2015). Modulating human sense of agency with non-invasive brain stimulation. *Cortex*, 69, 93–103. <http://doi.org/10.1016/j.cortex.2015.04.015>
- Kiani, R., Corthell, L., & Shadlen, M. N. (2014). Choice Certainty Is Informed by Both Evidence and Decision Time. *Neuron*, 84(6), 1329–1342. <http://doi.org/10.1016/j.neuron.2014.12.015>
- Kiefer, M., Sim, E.-J., Helbig, H., & Graf, M. (2011). Tracking the time course of action priming on object recognition: evidence for fast and slow influences of action on perception. *Journal of Cognitive Neuroscience*, 23(8), 1864–1874. <http://doi.org/10.1162/jocn.2010.21543>
- Kiesel, A., Wagener, A., Kunde, W., Hoffmann, J., Fallgatter, A. J., & Stöcker, C. (2006). Unconscious manipulation of free choice in humans. *Consciousness and Cognition*, 15(2), 397–408. <http://doi.org/10.1016/j.concog.2005.10.002>
- Klapp, S. T., & Haas, B. W. (2005). Nonconscious Influence of Masked Stimuli on Response Selection Is Limited to Concrete Stimulus-Response Associations. *Journal of Experimental Psychology: Human Perception and Performance*, 31(1), 193–209. <http://doi.org/10.1037/0096-1523.31.1.193>
- Klapp, S. T., & Hinkley, L. B. (2002). The negative compatibility effect: Unconscious inhibition influences reaction time and response selection. *Journal of Experimental Psychology: General*, 131(2), 255–269. <http://doi.org/10.1037/0096-3445.131.2.255>
- Kleiner, M., Brainard, D. H., & Pelli, D. G. (2007). What’s new in Psychtoolbox-3? *Perception*, 36(ECVP Abstract Supplement). <http://doi.org/10.1068/v070821>
- Knoblich, G., Elsner, B., Aschersleben, G., & Metzinger, T. (2003). Grounding the self in action. *Consciousness and Cognition*, 12(4), 487–494. <http://doi.org/10.1016/j.concog.2003.08.009>
- Kopp, B., Rist, F., & Mattler, U. (1996). N200 in the flanker task as a neurobehavioral

- tool for investigating executive control. *Psychophysiology*, 33(3), 282–294.
- Kordes-de Vaal, J. H. (1996). Intention and the omission bias: Omissions perceived as nondecisions. *Acta Psychologica*, 93(1–3), 161–172.  
[http://doi.org/10.1016/0001-6918\(96\)00027-3](http://doi.org/10.1016/0001-6918(96)00027-3)
- Koriat, A. (2000). The feeling of knowing: some metatheoretical implications for consciousness and control. *Consciousness and Cognition*, 9(2 Pt 1), 149–171.  
<http://doi.org/10.1006/ccog.2000.0433>
- Kühn, S., Brass, M., & Haggard, P. (2012). Feeling in control: Neural correlates of experience of agency. *Cortex*. <http://doi.org/10.1016/j.cortex.2012.09.002>
- Kühn, S., Nenchev, I., Haggard, P., Brass, M., Gallinat, J., & Voss, M. (2011). Whodunnit? Electrophysiological correlates of agency judgements. *PloS One*, 6(12), e28657. <http://doi.org/10.1371/journal.pone.0028657>
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2015). lmerTest: Tests in Linear Mixed Effects Models (Version 2.0-29). Retrieved from <https://cran.r-project.org/web/packages/lmerTest/index.html>
- Lafargue, G., & Franck, N. (2009). Effort awareness and sense of volition in schizophrenia. *Consciousness and Cognition*, 18(1), 277–289.  
<http://doi.org/10.1016/j.concog.2008.05.004>
- Larson, M. J., Clayson, P. E., & Clawson, A. (2014). Making sense of all the conflict: A theoretical review and critique of conflict-related ERPs. *International Journal of Psychophysiology*, 93(3), 283–297.  
<http://doi.org/10.1016/j.ijpsycho.2014.06.007>
- Limerick, H., Coyle, D., & Moore, J. W. (2014). The experience of agency in human-computer interactions: a review. *Frontiers in Human Neuroscience*, 8, 643.  
<http://doi.org/10.3389/fnhum.2014.00643>
- Lingnau, A., & Vorberg, D. (2005). The time course of response inhibition in masked priming. *Perception & Psychophysics*, 67(3), 545–557.
- Linser, K., & Goschke, T. (2007). Unconscious modulation of the conscious experience of voluntary control. *Cognition*, 104(3), 459–475.  
<http://doi.org/10.1016/j.cognition.2006.07.009>
- Li, P., Han, C., Lei, Y., Holroyd, C. B., & Li, H. (2011). Responsibility modulates neural mechanisms of outcome processing: An ERP study. *Psychophysiology*,

- 48(8), 1129–1133. <http://doi.org/10.1111/j.1469-8986.2011.01182.x>
- Li, P., Jia, S., Feng, T., Liu, Q., Suo, T., & Li, H. (2010). The influence of the diffusion of responsibility effect on outcome evaluations: Electrophysiological evidence from an ERP study. *NeuroImage*, 52(4), 1727–1733. <http://doi.org/10.1016/j.neuroimage.2010.04.275>
- Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review*, 1(4), 476–490. <http://doi.org/10.3758/BF03210951>
- Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: an open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience*, 8. <http://doi.org/10.3389/fnhum.2014.00213>
- Mansfield, K. L., van der Molen, M. W., Falkenstein, M., & van Boxtel, G. J. M. (2013). Temporal dynamics of interference in Simon and Eriksen tasks considered within the context of a dual-process model. *Brain and Cognition*, 82(3), 353–363. <http://doi.org/10.1016/j.bandc.2013.06.001>
- Marchetti, C., & Della Sala, S. (1998). Disentangling the Alien and Anarchic Hand. *Cognitive Neuropsychiatry*, 3(3), 191–207. <http://doi.org/10.1080/135468098396143>
- Matchock, R., & Mordkoff, J. T. (2007). Visual attention, reaction time, and self-reported alertness upon awakening from sleep bouts of varying lengths - Springer. *Exp Brain Res*, 178, 228–239. <http://doi.org/10.1007/s00221-006-0726-x>
- Mattler, U., & Palmer, S. (2012). Time course of free-choice priming effects explained by a simple accumulator model. *Cognition*, 123(3), 347–360. <http://doi.org/10.1016/j.cognition.2012.03.002>
- Meier, B. P., & Hinsz, V. B. (2004). A comparison of human aggression committed by groups and individuals: An interindividual–intergroup discontinuity. *Journal of Experimental Social Psychology*, 40(4), 551–559. <http://doi.org/10.1016/j.jesp.2003.11.002>
- Metcalfe, J., Butterfield, B., Habeck, C., & Stern, Y. (2012). Neural Correlates of People's Hypercorrection of Their False Beliefs. *Journal of Cognitive Neuroscience*. [http://doi.org/10.1162/jocn\\_a\\_00228](http://doi.org/10.1162/jocn_a_00228)

- Metcalfe, J., Eich, T. S., & Castel, A. D. (2010). Metacognition of agency across the lifespan. *Cognition*, 116(2), 267–282.  
<http://doi.org/10.1016/j.cognition.2010.05.009>
- Metcalfe, J., Eich, T. S., & Miele, D. B. (2013). Metacognition of agency: proximal action and distal outcome. *Experimental Brain Research*, 229(3), 485–496.  
<http://doi.org/10.1007/s00221-012-3371-6>
- Metcalfe, J., & Greene, M. J. (2007). Metacognition of agency. *Journal of Experimental Psychology. General*, 136(2), 184–199.  
<http://doi.org/10.1037/0096-3445.136.2.184>
- Metcalfe, J., Van Snellenberg, J. X., DeRosse, P., Balsam, P., & Malhotra, A. K. (2012). Judgements of agency in schizophrenia: an impairment in autonoetic metacognition. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 367(1594), 1391–1400.  
<http://doi.org/10.1098/rstb.2012.0006>
- Miele, D. B., Wager, T. D., Mitchell, J. P., & Metcalfe, J. (2011). Dissociating neural correlates of action monitoring and metacognition of agency. *Journal of Cognitive Neuroscience*, 23(11), 3620–3636.  
[http://doi.org/10.1162/jocn\\_a\\_00052](http://doi.org/10.1162/jocn_a_00052)
- Mill, R. D., & O'Connor, A. R. (2014). Question format shifts bias away from the emphasised response in tests of recognition memory. *Consciousness and Cognition*, 30, 91–104. <http://doi.org/10.1016/j.concog.2014.09.006>
- Moore, J. W., & Fletcher, P. C. (2012). Sense of agency in health and disease: A review of cue integration approaches. *Consciousness and Cognition*, 21(1), 59–68. <http://doi.org/10.1016/j.concog.2011.08.010>
- Moore, J. W., & Haggard, P. (2008). Awareness of action: Inference and prediction. *Consciousness and Cognition*, 17(1), 136–144.  
<http://doi.org/10.1016/j.concog.2006.12.004>
- Moore, J. W., Middleton, D., Haggard, P., & Fletcher, P. C. (2012). Exploring implicit and explicit aspects of sense of agency. *Consciousness and Cognition*, 21(4), 1748–1753. <http://doi.org/10.1016/j.concog.2012.10.005>
- Moore, J. W., & Obhi, S. S. (2012). Intentional binding and the sense of agency: a review. *Consciousness and Cognition*, 21(1), 546–561.

- <http://doi.org/10.1016/j.concog.2011.12.002>
- Moore, J. W., Ruge, D., Wenke, D., Rothwell, J., & Haggard, P. (2010). Disrupting the experience of control in the human brain: pre-supplementary motor area contributes to the sense of agency. *Proceedings. Biological Sciences / The Royal Society*, 277(1693), 2503–2509. <http://doi.org/10.1098/rspb.2010.0404>
- Moretto, G., Walsh, E., & Haggard, P. (2011). Experience of agency and sense of responsibility. *Consciousness and Cognition*.  
<http://doi.org/10.1016/j.concog.2011.08.014>
- Morsella, E., Wilson, L. E., Berger, C. C., Honhongva, M., Gazzaley, A., & Bargh, J. A. (2009). Subjective aspects of cognitive control at different stages of processing. *Attention, Perception & Psychophysics*, 71(8), 1807–1824.  
<http://doi.org/10.3758/APP.71.8.1807>
- Musseler, J., & Wuhr, P. (2002). Response-evoked interference in visual encoding. In W. Prinz & B. Hommel (Eds.), *Common Mechanisms in Perception and Action: Attention and Performance Volume XIX: Attention and Performance v. 19* (pp. 520–537). OUP Oxford.
- Mynatt, C., & Sherman, S. J. (1975). Responsibility attribution in groups and individuals: A direct test of the diffusion of responsibility hypothesis. *Journal of Personality and Social Psychology*, 32(6), 1111.
- Nahab, F. B., Kundu, P., Gallea, C., Kakareka, J., Pursley, R., Pohida, T., ... Hallett, M. (2011). The neural processes underlying self-agency. *Cerebral Cortex*, 21(1), 48–55. <http://doi.org/10.1093/cercor/bhq059>
- Nahmias, E. (2005). Agency, authorship, and illusion. *Consciousness and Cognition*, 14(4), 771–785. <http://doi.org/10.1016/j.concog.2005.07.002>
- Nelson, T. O., & Narrens, L. (1990). Metamemory: A Theoretical Framework and New Findings. *Psychology of Learning and Motivation*, 26, 125–173.
- Nikolaev, A. R., Ziessler, M., Dimova, K., & van Leeuwen, C. (2008). Anticipated action consequences as a nexus between action and perception: evidence from event-related potentials. *Biological Psychology*, 78(1), 53–65.  
<http://doi.org/10.1016/j.biopsycho.2007.12.010>
- O'Connor, P. A., & Neill, W. T. (2011). Does subliminal priming of free response choices depend on task set or automatic response activation? *Consciousness*



- and Cognition*, 20(2), 280–287. <http://doi.org/10.1016/j.concog.2010.08.007>
- Oliveira, F. T. P., McDonald, J. J., & Goodman, D. (2007). Performance Monitoring in the Anterior Cingulate is Not All Error Related: Expectancy Deviation and the Representation of Action-Outcome Associations. *Journal of Cognitive Neuroscience*, 19(12), 1994–2004. <http://doi.org/10.1162/jocn.2007.19.12.1994>
- Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2011). FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. *Computational Intelligence and Neuroscience*, 2011. <http://doi.org/10.1155/2011/156869>
- Pacherie, E. (2007). The Sense of Control and the Sense of Agency. *Psyche*, 13(1), 1–30.
- Pacherie, E. (2008). The phenomenology of action: a conceptual framework. *Cognition*, 107(1), 179–217. <http://doi.org/10.1016/j.cognition.2007.09.003>
- Padrao, G., Gonzalez-Franco, M., Sanchez-Vives, M. V., Slater, M., & Rodriguez-Fornells, A. (2016). Violating body movement semantics: Neural signatures of self-generated and external-generated errors. *NeuroImage*, 124, Part A, 147–156. <http://doi.org/10.1016/j.neuroimage.2015.08.022>
- Pailing, P. E., & Segalowitz, S. J. (2004). The effects of uncertainty in error monitoring on associated ERPs. *Brain and Cognition*, 56(2), 215–233. <http://doi.org/10.1016/j.bandc.2004.06.005>
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision*, 10(4), 437–442.
- Pfister, R., & Janczyk, M. (2013). Confidence intervals for two sample means: Calculation, interpretation, and a few simple rules. *Advances in Cognitive Psychology*, 9(2), 74–80. <http://doi.org/10.2478/v10053-008-0133-x>
- Proust, J. (2008). XIII—Epistemic Agency and Metacognition: An Externalist View. *Proceedings of the Aristotelian Society (Hardback)*, 108(1pt3), 241–268. <http://doi.org/10.1111/j.1467-9264.2008.00245.x>
- R Core Team. (2015). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria. Retrieved from <http://www.R-project.org>
- Reber, R., Winkielman, P., & Schwarz, N. (1998). Effects of Perceptual Fluency on

- Affective Judgments. *Psychological Science*, 9(1), 45–48.  
<http://doi.org/10.1111/1467-9280.00008>
- Resulaj, A., Kiani, R., Wolpert, D. M., & Shadlen, M. N. (2009). Changes of mind in decision-making. *Nature*, 461(7261), 263–266.  
<http://doi.org/10.1038/nature08275>
- Reznik, D., Henkin, Y., Levy, O., & Mukamel, R. (2015). Perceived Loudness of Self-Generated Sounds Is Differentially Modified by Expected Sound Intensity. *PLoS ONE*, 10(5), e0127651. <http://doi.org/10.1371/journal.pone.0127651>
- Ritterband-Rosenbaum, A., Karabanov, A. N., Christensen, M. S., & Nielsen, J. B. (2014). 10 Hz rTMS over right parietal cortex alters sense of agency during self-controlled movements. *Frontiers in Human Neuroscience*, 8, 471.  
<http://doi.org/10.3389/fnhum.2014.00471>
- Rodríguez-Fornells, A., Kurzbuch, A. R., & Münte, T. F. (2002). Time Course of Error Detection and Correction in Humans: Neurophysiological Evidence. *The Journal of Neuroscience*, 22(22), 9990–9996.
- Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs: General and Applied*, 80(1), 1.  
<http://doi.org/10.1037/h0092976>
- Saito, N., Takahata, K., Murai, T., & Takahashi, H. (2015). Discrepancy between explicit judgement of agency and implicit feeling of agency: Implications for sense of agency and its disorders. *Consciousness and Cognition*, 37, 1–7.  
<http://doi.org/10.1016/j.concog.2015.07.011>
- San Martín, R. (2012). Event-related potential studies of outcome processing and feedback-guided learning. *Frontiers in Human Neuroscience*, 6, 304.  
<http://doi.org/10.3389/fnhum.2012.00304>
- Sato, A. (2009). Both motor prediction and conceptual congruency between preview and action-effect contribute to explicit judgment of agency. *Cognition*, 110(1), 74–83. <http://doi.org/10.1016/j.cognition.2008.10.011>
- Sato, A., & Yasuda, A. (2005). Illusion of sense of self-agency: discrepancy between the predicted and actual sensory consequences of actions modulates the sense of self-agency, but not the sense of self-ownership. *Cognition*, 94(3), 241–255.  
<http://doi.org/10.1016/j.cognition.2004.04.003>

- Scheffers, M. K., & Coles, M. G. H. (2000). Performance Monitoring in a Confusing World: Error-Related Brain Activity, Judgments of Response Accuracy, and Types of Errors. *Journal of Experimental Psychology*, 26(1), 141–151.
- Schel, M. A., Brass, M., Haggard, P., Ridderinkhof, K. R., & Crone, E. A. (2014). Neural correlates of intentional and stimulus-driven inhibition: a comparison. *Frontiers in Human Neuroscience*, 8, 27.  
<http://doi.org/10.3389/fnhum.2014.00027>
- Schiffer, A.-M., Waszak, F., & Yeung, N. (2015). The role of prediction and outcomes in adaptive cognitive control. *Journal of Physiology-Paris*, 109(1–3), 38–52.  
<http://doi.org/10.1016/j.jphysparis.2015.02.001>
- Schlaghecken, F., & Eimer, M. (2004). Masked prime stimuli can bias “free” choices between response alternatives. *Psychonomic Bulletin & Review*, 11(3), 463–468. <http://doi.org/10.3758/BF03196596>
- Schubö, A., Prinz, W., & Aschersleben, G. (2004). Perceiving while acting: action affects perception. *Psychological Research*, 68(4), 208–215.  
<http://doi.org/10.1007/s00426-003-0133-x>
- Schütz-Bosbach, S., & Prinz, W. (2007). Perceptual resonance: action-induced modulation of perception. *Trends in Cognitive Sciences*, 11(8), 349–355.  
<http://doi.org/10.1016/j.tics.2007.06.005>
- Sebanz, N., & Lackner, U. (2007). Who’s calling the shots? Intentional content and feelings of control. *Consciousness and Cognition*, 16(4), 859–876.  
<http://doi.org/10.1016/j.concog.2006.08.002>
- Shergill, S. S., Samson, G., Bays, P. M., Frith, C. D., & Wolpert, D. M. (2005). Evidence for Sensory Prediction Deficits in Schizophrenia. *American Journal of Psychiatry*, 162(12), 2384–2386. <http://doi.org/10.1176/appi.ajp.162.12.2384>
- Sidarus, N., Chambon, V., & Haggard, P. (2013). Priming of actions increases sense of control over unexpected outcomes. *Consciousness and Cognition*, 22(4), 1403–1411. <http://doi.org/10.1016/j.concog.2013.09.008>
- Simon-Thomas, E. R., & Knight, R. T. (2005). Affective and cognitive modulation of performance monitoring: Behavioral and ERP evidence. *Cognitive, Affective, & Behavioral Neuroscience*, 5(3), 362–372. <http://doi.org/10.3758/CABN.5.3.362>
- Snijders, T. A. B., & Bosker, R. J. (1999). *Multilevel Analysis: An Introduction to*

*Basic and Advanced Multilevel Modeling*. SAGE.

- Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General*, 117(1), 34–50. <http://doi.org/10.1037/0096-3445.117.1.34>
- Spence, S. (2009). *The actor's brain: Exploring the cognitive neuroscience of free will* (1st ed.). OUP Oxford.
- Spengler, S., von Cramon, D. Y., & Brass, M. (2009). Was it me or was it you? How the sense of agency originates from ideomotor learning revealed by fMRI. *NeuroImage*, 46(1), 290–298. <http://doi.org/10.1016/j.neuroimage.2009.01.047>
- Sperduti, M., Delaveau, P., Fossati, P., & Nadel, J. (2011). Different brain structures related to self- and external-agency attribution: a brief review and meta-analysis. *Brain Structure & Function*, 216(2), 151–157. <http://doi.org/10.1007/s00429-010-0298-1>
- Steinhauser, M., & Kiesel, A. (2011). Performance monitoring and the causal attribution of errors. *Cognitive, Affective & Behavioral Neuroscience*, 11(3), 309–320. <http://doi.org/10.3758/s13415-011-0033-2>
- Stenner, M.-P., Bauer, M., Heinze, H.-J., Haggard, P., & Dolan, R. J. (2014). Parallel processing streams for motor output and sensory prediction during action preparation. *Journal of Neurophysiology*, jn.00616.2014. <http://doi.org/10.1152/jn.00616.2014>
- Stenner, M.-P., Bauer, M., Sidarus, N., Heinze, H.-J., Haggard, P., & Dolan, R. J. (2014). Subliminal action priming modulates the perceived intensity of sensory action consequences. *Cognition*, 130(2), 227–235. <http://doi.org/10.1016/j.cognition.2013.11.008>
- Sumner, P., Nachev, P., Morris, P., Peters, A. M., Jackson, S. R., Kennard, C., & Husain, M. (2007). Human Medial Frontal Cortex Mediates Unconscious Inhibition of Voluntary Action. *Neuron*, 54(5), 697–711. <http://doi.org/10.1016/j.neuron.2007.05.016>
- Synofzik, M., Vosgerau, G., & Newen, A. (2008). Beyond the comparator model: a multifactorial two-step account of agency. *Consciousness and Cognition*, 17(1), 219–239. <http://doi.org/10.1016/j.concog.2007.03.010>
- Synofzik, M., Vosgerau, G., & Voss, M. (2013). The experience of agency: an

- interplay between prediction and postdiction. *Frontiers in Psychology*, 4. <http://doi.org/10.3389/fpsyg.2013.00127>
- Takahata, K., Takahashi, H., Maeda, T., Umeda, S., Suhara, T., Mimura, M., & Kato, M. (2012). It's Not My Fault: Postdictive Modulation of Intentional Binding by Monetary Gains and Losses. *PLoS ONE*, 7(12), e53421. <http://doi.org/10.1371/journal.pone.0053421>
- Taylor, D. A. (1977). Time course of context effects. *Journal of Experimental Psychology. General*, 106(4), 404–426. <http://doi.org/10.1037/0096-3445.106.4.404>
- Thorndike, E. L., & Bruce, D. (1911). *Animal Intelligence: Experimental Studies*. Transaction Publishers.
- Tibon, R., & Levy, D. A. (2015). Striking a balance: analyzing unbalanced event-related potential data. *Quantitative Psychology and Measurement*, 6, 555. <http://doi.org/10.3389/fpsyg.2015.00555>
- Timm, J., SanMiguel, I., Saupe, K., & Schröger, E. (2013). The N1-suppression effect for self-initiated sounds is independent of attention. *BMC Neuroscience*, 14(1), 2. <http://doi.org/10.1186/1471-2202-14-2>
- Tsakiris, M., Haggard, P., Franck, N., Mainy, N., & Sirigu, A. (2005). A specific role for efferent information in self-recognition. *Cognition*, 96(3), 215–231. <http://doi.org/10.1016/j.cognition.2004.08.002>
- Tsakiris, M., Longo, M. R., & Haggard, P. (2010). Having a body versus moving your body: Neural signatures of agency and body-ownership. *Neuropsychologia*, 48(9), 2740–2749. <http://doi.org/10.1016/j.neuropsychologia.2010.05.021>
- van Gaal, S., Lamme, V. A. F., & Ridderinkhof, K. R. (2010). Unconsciously Triggered Conflict Adaptation. *PLoS ONE*, 5(7), e11508. <http://doi.org/10.1371/journal.pone.0011508>
- van Gaal, S., Scholte, H. S., Lamme, V. A. F., Fahrenfort, J. J., & Ridderinkhof, K. R. (2010). Pre-SMA graymatter density predicts individual differences in action selection in the face of conscious and unconscious response conflict. *Journal of Cognitive Neuroscience*, 23(2), 382–390. <http://doi.org/10.1162/jocn.2010.21444>
- van Veen, V., & Carter, C. S. (2002). The Timing of Action-Monitoring Processes in

- the Anterior Cingulate Cortex. *Journal of Cognitive Neuroscience*, 14(4), 593–602. <http://doi.org/10.1162/08989290260045837>
- Verleger, R., & Jaskowski, P. (2008). Disentangling neural processing of masked and masking stimulus by means of event-related contralateral - ipsilateral differences of EEG potentials. *Advances in Cognitive Psychology*, 3(1-2), 193–210. <http://doi.org/10.2478/v10053-008-0025-0>
- Vermeiren, A., & Cleeremans, A. (2012). The validity of d' measures. *PloS One*, 7(2), e31595. <http://doi.org/10.1371/journal.pone.0031595>
- Vogel, E. K., & Luck, S. J. (2000). The visual N1 component as an index of a discrimination process. *Psychophysiology*, 37(2), 190–203.
- Vorberg, D., Mattler, U., Heinecke, A., Schmidt, T., & Schwarzbach, J. (2003). Different time courses for visual perception and action priming. *Proceedings of the National Academy of Sciences*, 100(10), 6275–6280. <http://doi.org/10.1073/pnas.0931489100>
- Voss, M., Moore, J. W., Hauser, M., Gallinat, J., Heinz, A., & Haggard, P. (2010). Altered awareness of action in schizophrenia: a specific deficit in predicting action consequences. *Brain*, 133(10), 3104–3112. <http://doi.org/10.1093/brain/awq152>
- Wagenmakers, E.-J. (2007). A practical solution to the pervasive problems of p values. *Psychonomic Bulletin & Review*, 14(5), 779–804. <http://doi.org/10.3758/BF03194105>
- Wallach, M. A., Kogan, N., & Bem, D. J. (1964). Diffusion of responsibility and level of risk taking in groups. *The Journal of Abnormal and Social Psychology*, 68(3), 263.
- Wang, B., Xiang, L., & Li, J. (2013). Does conflict control occur without awareness? Evidence from an ERP study. *Brain Research*, 1490, 161–169. <http://doi.org/10.1016/j.brainres.2012.10.053>
- Wascher, E., Reinhard, M., Wauschkuhn, B., & Verleger, R. (1999). Spatial S-R Compatibility with Centrally Presented Stimuli: An Event-Related Asymmetry Study on Dimensional Overlap. *Journal of Cognitive Neuroscience*, 11(2), 214–229. <http://doi.org/10.1162/089892999563346>
- Wegner, D. M. (2004). Précis of the illusion of conscious will. *The Behavioral and*

- Brain Sciences*, 27(5), 649–659; discussion 659–692.
- Wegner, D. M., & Wheatley, T. (1999). Apparent mental causation. Sources of the experience of will. *The American Psychologist*, 54(7), 480–492.
- Weiss, C., Herwig, A., & Schütz-Bosbach, S. (2011). The self in action effects: Selective attenuation of self-generated sounds. *Cognition*, 121(2), 207–218. <http://doi.org/10.1016/j.cognition.2011.06.011>
- Weiss, C., Tsakiris, M., Haggard, P., & Schütz-Bosbach, S. (2013). Agency in the sensorimotor system and its relation to explicit action awareness. *Neuropsychologia*. <http://doi.org/10.1016/j.neuropsychologia.2013.09.034>
- Wenke, D., Fleming, S. M., & Haggard, P. (2010). Subliminal priming of actions influences sense of control over effects of action. *Cognition*, 115(1), 26–38. <http://doi.org/10.1016/j.cognition.2009.10.016>
- Wenke, D., Gaschler, R., Nattkemper, D., & Frensch, P. A. (2009). Strategic influences on implementing instructions for future actions. *Psychological Research*, 73(4), 587–601. <http://doi.org/10.1007/s00426-009-0239-x>
- Whittlesea, B. W. A., & Leboe, J. P. (2003). Two fluency heuristics (and how to tell them apart). *Journal of Memory and Language*, 49(1), 62–79. [http://doi.org/10.1016/S0749-596X\(03\)00009-3](http://doi.org/10.1016/S0749-596X(03)00009-3)
- Willemsen, R., Hoormann, J., Hohnsbein, J., & Falkenstein, M. (2004). Central and parietal event-related lateralizations in a flanker task. *Psychophysiology*, 41(5), 762–771.
- Winkelman, P., Ziembowicz, M., & Nowak, A. (2015). The coherent and fluent mind: how unified consciousness is constructed from cross-modal inputs via integrated processing experiences. *Consciousness Research*, 6, 83. <http://doi.org/10.3389/fpsyg.2015.00083>
- Wolpe, N., Haggard, P., Siebner, H. R., & Rowe, J. B. (2013). Cue integration and the perception of action in intentional binding. *Experimental Brain Research. Experimentelle Hirnforschung. Experimentation Cerebrale*. <http://doi.org/10.1007/s00221-013-3419-2>
- Wolpert, D. M., Ghahramani, Z., & Jordan, M. I. (1995). An internal model for sensorimotor integration. *Science (New York, N.Y.)*, 269(5232), 1880–1882.

- Yeung, N., Botvinick, M. M., & Cohen, J. D. (2004). The neural basis of error detection: conflict monitoring and the error-related negativity. *Psychological Review*, 111(4), 931–959. <http://doi.org/10.1037/0033-295X.111.4.939>
- Yeung, N., Holroyd, C. B., & Cohen, J. D. (2005). ERP Correlates of Feedback and Reward Processing in the Presence and Absence of Response Choice. *Cerebral Cortex*, 15(5), 535–544. <http://doi.org/10.1093/cercor/bhh153>
- Yeung, N., & Nieuwenhuis, S. (2009). Dissociating Response Conflict and Error Likelihood in Anterior Cingulate Cortex. *The Journal of Neuroscience*, 29(46), 14506–14510. <http://doi.org/10.1523/JNEUROSCI.3615-09.2009>
- Yeung, N., & Summerfield, C. (2012). Metacognition in human decision-making: confidence and error monitoring. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1594), 1310–1321. <http://doi.org/10.1098/rstb.2011.0416>
- Yeung, N., & Summerfield, C. (2014). Shared Mechanisms for Confidence Judgements and Error Detection in Human Decision Making. In S. M. Fleming & C. D. Frith (Eds.), *The Cognitive Neuroscience of Metacognition* (pp. 147–167). Springer Berlin Heidelberg. Retrieved from [http://link.springer.com/chapter/10.1007/978-3-642-45190-4\\_7](http://link.springer.com/chapter/10.1007/978-3-642-45190-4_7)
- Yomogida, Y., Sugiura, M., Sassa, Y., Wakusawa, K., Sekiguchi, A., Fukushima, A., ... Kawashima, R. (2010). The neural basis of agency: an fMRI study. *NeuroImage*, 50(1), 198–207. <http://doi.org/10.1016/j.neuroimage.2009.12.054>
- Yoshie, M., & Haggard, P. (2013). Negative Emotional Outcomes Attenuate Sense of Agency over Voluntary Actions. *Current Biology*. <http://doi.org/10.1016/j.cub.2013.08.034>
- Yu, A. J., & Dayan, P. (2005). Uncertainty, neuromodulation, and attention. *Neuron*, 46(4), 681–692. <http://doi.org/10.1016/j.neuron.2005.04.026>
- Zalla, T., Miele, D., Leboyer, M., & Metcalfe, J. (2015). Metacognition of agency and theory of mind in adults with high functioning autism. *Consciousness and Cognition*, 31, 126–138. <http://doi.org/10.1016/j.concog.2014.11.001>